

From: Fredman, Jeffrey
Sent: Monday, October 01, 2001 3:42 PM
To: STIC-Biotech/ChemLib
Cc: Lacourciere, Karen
Subject: FW: Sequence search approval 09/757100

I Approve.

Jeff Fredman

-----Original Message-----

Fr m: Lacourciere, Karen
Sent: Monday, October 01, 2001 3:06 PM
T : Fredman, Jeffrey
Subject: Sequence search approval 09/757100

Jeff, could you approve this sequence search? Each sequence is ~20 nt's long and they are antisense targeted to one sequence. Thank-you!
Karen

Please search the following SEQ ID NO:'s for 09/757,100 in both the commercial databases and the pending files (interference)
3, 4, 6-12, 14-18, 20, 23, 30, 31 and 33.

Please limit the length of oligos to less than 50 nucleotides long.
Thank-you!

Karen A. Lacourciere Ph.D.
✓ CM1 11D09 GAU 1635 ✓
(703) 308-7523

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 1E01 TEL: 308-3534

Point of Contact:
Toby Port
Technical Info. Specialist
CMI 1E01 TEL: 308-3234

P

SEARCH REQUEST FORM

Access DB#

52048

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>TOBY PORT</u>	NA Sequence (#) <u>19</u>	STN _____
Searcher Phone #: <u>308-3534</u>	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbi: _____
Date Searcher Picked Up: <u>10/2</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>10/3</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: <u>15</u>	Fulltext _____	Sequence Systems <u>CS</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: <u>10</u>	Other _____	Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:50 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-18

Perfect score: 20

Sequence: 1 tttaaccagatggtcattc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba1.*

2: gb_ba2.*

3: gb_ba3.*

4: gb_in1.*

5: gb_in2.*

6: gb_in3.*

7: gb_om.*

8: gb_ov.*

9: gb_pat1.*

10: gb_pat2.*

11: gb_ph.*

12: gb_pl1.*

13: gb_pl2.*

14: gb_pl3.*

15: gb_pl4.*

16: em_ba1.*

17: em_ba2.*

18: em_fun.*

19: em_htgo_hum.*

20: em_htgo_inv.*

21: em_htgo_rod.*

22: em_htg_hum1.*

23: em_htg_hum2.*

24: em_htg_hum3.*

25: em_htg_hum4.*

26: em_htg_hum5.*

27: em_htg_hum6.*

28: em_htg_hum7.*

29: em_htg_hum8.*

30: em_htg_inv1.*

31: em_htg_inv2.*

32: em_htg_other.*

33: em_htg_rod.*

34: em_hum1.*

35: em_hum2.*

36: em_hum3.*

37: em_hum4.*

38: em_hum5.*

39: em_hum6.*

40: em_hum7.*

41: em_in.*

42: em_om.*

43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vil2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_rol.*
95: gb_ro2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
c 1	12.8	64.0	20	9 A10280	A10280 oligonucleo
2	12.8	64.0	29	10 I82248	I82248 Sequence 32
3	12.8	64.0	29	10 I90782	I90782 Sequence 32
c 4	12.8	64.0	45	97 HUMKRT10AG	L20218 Human Kerat
c 5	12.6	63.0	35	10 I64891	I64891 Sequence 7
c 6	12.6	63.0	41	10 E13249	E13249 Oligonucleo
7	12.4	62.0	30	9 AR004965	AR004965 Sequence
8	12.4	62.0	30	10 I40158	I40158 Sequence 4

```

9 12.2 61.0 20 9 AR016116 Sequence
10 12.2 61.0 20 9 AR019114 Sequence
11 12.2 61.0 28 9 AR090906 Sequence
12 12.2 61.0 33 9 AR034456 Sequence
13 12.2 61.0 33 9 AR070401 Sequence
14 12.2 61.0 34 9 AR026654 Sequence
15 12.2 61.0 29 10 AX067989 Sequence
16 12.2 60.0 33 10 AX076373 Sequence
17 12.2 60.0 34 9 A03829 Sequence
18 12.2 60.0 40 9 AR070497 Sequence
19 12.2 60.0 44 10 E14337 Sequence
20 11.8 59.0 16 9 A42578 Sequence
21 11.8 59.0 16 9 A42579 Sequence
22 11.8 59.0 16 9 A88767 Sequence
23 11.8 59.0 16 9 A88768 Sequence
24 11.8 59.0 23 9 A04127 Sequence
25 11.8 59.0 25 9 A067151 Sequence
26 11.8 59.0 35 9 AR075963 Sequence
27 11.8 59.0 38 9 A67648 Sequence
28 11.8 59.0 38 9 AR089824 Sequence
29 11.6 58.0 21 54 HUM66RVA D50147 A PCR prime
30 11.6 58.0 23 9 A37961 Sequence
31 11.6 58.0 23 10 I21766 Sequence
32 11.6 58.0 32 9 AR011102 Sequence
33 11.6 58.0 32 9 AR038261 Sequence
34 11.6 58.0 32 9 AR075191 Sequence
35 11.6 58.0 32 10 I58550 Sequence
36 11.6 58.0 32 10 I61215 Sequence
37 11.6 58.0 33 9 AR054052 Sequence
38 11.6 58.0 38 9 AR011107 Sequence
39 11.6 58.0 38 9 AR038266 Sequence
40 11.6 58.0 38 9 AR075196 Sequence
41 11.6 58.0 38 10 I58555 Sequence
42 11.6 58.0 38 10 I61220 Sequence
43 11.6 58.0 49 10 E12236 Synthetic D
44 11.6 58.0 49 10 E12238 Synthetic D
45 11.6 58.0 50 10 E22340 DNA encodin

```

ALIGNMENTS

```

RESULT 1
A10280/c A10280 20 bp DNA PAT 25-AUG-1993
LOCUS oligonucleotide.
DEFINITION A10280
ACCESSION A10280
VERSION A10280.1 GI:413528
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS
JOURNAL
FEATURES
Patent: DE 3802040-A 1 03-AUG-1989;
Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 8 a 1 c 6 g 5 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 9; Length 20;
Best Local Similarity 87.5%; Pred. No. 6.3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttcaaccagatggc 16
Db 20 TTTCACCATGCTTC 5

RESULT 2
A10280/c A10280 20 bp DNA PAT 25-AUG-1993
LOCUS oligonucleotide.
DEFINITION A10280
ACCESSION A10280
VERSION A10280.1 GI:413528
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS
JOURNAL
FEATURES
Patent: DE 3802040-A 1 03-AUG-1989;
Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 8 a 1 c 6 g 5 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 9; Length 20;
Best Local Similarity 87.5%; Pred. No. 6.3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttcaaccagatggc 16
Db 20 TTTCACCATGCTTC 5

```

```

I82248 I82248 29 bp DNA PAT 10-JUN-1998
LOCUS Sequence 32 from patent US 5712118.
DEFINITION I82248
ACCESSION I82248
VERSION I82248.1 GI:3210545
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 29)
AUTHORS Murphy, R.F.
TITLE Vaccine for brachamella catarrhalis
JOURNAL Patent: US 5712118-A 32 27-JAN-1998;
FEATURES Location/Qualifiers
1..29
Source /organism="unknown"
BASE COUNT 11 a 5 c 5 g 8 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 10; Length 29;
Best Local Similarity 87.5%; Pred. No. 6.3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ttcaaccagatggc 17
Db 10 TTTCATCCAGATGATCA 25

RESULT 3
I82248 I82248 29 bp DNA PAT 01-DEC-1998
LOCUS Sequence 32 from patent US 5725862.
DEFINITION I82248
ACCESSION I82248
VERSION I82248.1 GI:3935252
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 29)
AUTHORS Murphy, R.F.
TITLE Vaccine for brachamella catarrhalis
JOURNAL Patent: US 5725862-A 32 10-MAR-1998;
FEATURES Location/Qualifiers
1..29
Source /organism="unknown"
BASE COUNT 11 a 5 c 5 g 8 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 10; Length 29;
Best Local Similarity 87.5%; Pred. No. 6.3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ttcaaccagatggc 17
Db 10 TTTCATCCAGATGATCA 25

RESULT 4
HUMKRT10AG/c HUMKRT10AG 45 bp DNA PRI 06-JAN-1995
LOCUS Human keratin 10 (KRT10) gene, partial cds including polymorphism.
DEFINITION L20218
ACCESSION L20218
VERSION L20218.1 GI:307087
KEYWORDS keratin; keratin 10; polymorphism; suprabasal keratin.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 45)
AUTHORS Rothnagel, J.J., Dominey, A., Fisher, M., Axtell, S., Pittelkow, M.,
Anton-Lamprecht, I., Hohl, D. and Roop, D.

```

TITLE Identification of mutational hot spots in the suprabasal keratin genes from patients with epidermolytic hyperkeratosis
JOURNAL Unpublished (1993)
FEATURES Location/Qualifiers
source 1..45
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="Lymphocyte"
/dev_stage="Adult"
/tissue_type="Blood"
/map="17q21-q23"
/note="from EHK patient"
1..45
/gene="KRT10"
CDS <1..>45
/gene="KRT10"
/note="helix-initiation motif of 1A segment of rod domain"
/codon_start=1
/function="intermediate filament precursor"
/db_xref="GDB:G00-118-828"
/product="keratin 10"
/protein_id="AAB59438.1"
/db_xref="GI:307088"
/translation="KVTQMQLNDHLASYL"
29
variation
/gene="KRT10"
/note="polymorphism results in an Arg->His mutation at amino acid position #10 in the HK10 rod; G00-118-828"
/phenotype="wild-type (g)"
/replace="g"
14 a 12 c 8 g 11 t
BASE COUNT
ORIGIN
Query Match 64.0%; Score 12.8; DB 97; Length 45;
Best Local Similarity 87.5%; Pred. No. 6.2e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 aaccagatgggtcattc 20
||||| |||||
Db 36 ACCCAGGTGGTCATTC 21
RESULT 5
LOCUS I64891 35 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 7 from patent US 5665874.
ACCESSION I64891
VERSION I64891.1 GI:2481785
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 35)
AUTHORS Kuhajda,F.P. and Pasternack,G.R.
TITLE Cancer related antigen
JOURNAL Patent: US 5665874-A 7 09-SEP-1997;
FEATURES Location/Qualifiers
source 1..35
/organism="unknown"
BASE COUNT 9 a 12 c 6 g 8 t
ORIGIN
Query Match 63.0%; Score 12.6; DB 10; Length 35;
Best Local Similarity 78.9%; Pred. No. 7.9e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 2 ttcaaccagatgggtcattc 20
||||| ||||| |||||
Db 35 TTCAAGAAGATGGCCATGC 17

RESULT 6
LOCUS E13249/c 41 bp DNA PAT 24-JUN-1998
DEFINITION Oligonucleotide for microgene construction.
ACCESSION E13249
VERSION E13249.1 GI:3252054
KEYWORDS JP 1997154585-A/9.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Shiba,K..
TITLE FORMATION OF RANDOM POLYMER OF MICROGENE
JOURNAL Patent: JP 1997154585-A 9 17-JUN-1997;
COMMENT KAGAKU GIJUTSU SHINKO JIGYODAN
OS None
OC Artificial sequences.
PN JP 1997154585-A/9
PD 17-JUN-1997
PF 06-DEC-1995 JP 1995318396
PI SHIBA KIYOTAKA
PC C12N15/09,C07H21/04//C12N9/00;
CC strandedness: Single;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..41
/organism="Artificial sequences".
FT Location/Qualifiers
1..41
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 10 a 10 c 15 g 6 t
ORIGIN
Query Match 63.0%; Score 12.6; DB 10; Length 41;
Best Local Similarity 78.9%; Pred. No. 7.9e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 tttaaccagatgggtcatt 19
||||| ||||| |||||
Db 32 TTTCACCATGATCCCCCTT 14
RESULT 7
LOCUS AR004965 30 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 4 from patent US 5747312.
ACCESSION AR004965
VERSION AR004965.1 GI:3965844
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Wei,Y. and Sutton,G.G. III.
TITLE Human ALKB polypeptide
JOURNAL Patent: US 5747312-A 4 05-MAY-1998;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
BASE COUNT 9 a 9 c 7 g 5 t
ORIGIN
Query Match 62.0%; Score 12.4; DB 9; Length 30;
Best Local Similarity 92.9%; Pred. No. 1e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 tttaaccagatggg 14
||||| |||||
Db 9 TTTCATCCAGATGG 22

RESULT 8
 LOCUS I40158 30 bp DNA PAT 13-MAY-1997
 DEFINITION Sequence 4 from patent US 5618717.
 ACCESSION I40158
 VERSION I40158.1 GI:2083163
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 30)
 AUTHORS Wei,Y. and Sutton,G.G. III.
 TITLE DNA encoding human ALKB
 JOURNAL Patent: US 5618717-A 4 08-APR-1997;
 FEATURES Location/Qualifiers
 source 1..30
 BASE COUNT 9 a 9 c 7 g 5 t
 ORIGIN
 Query Match 62.0%; Score 12.4; DB 10; Length 30;
 Best Local Similarity 92.9%; Pred. No. 1e+05;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 tttaaccagatgg 14
 ||||| |||||
 Db 9 TTTCATCCAGATGG 22
 ||||| |||||
 RESULT 9
 LOCUS AR016116 20 bp DNA PAT 05-DEC-1998
 DEFINITION Sequence 4 from patent US 5776682.
 ACCESSION AR016116
 VERSION AR016116.1 GI:3972393
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 20)
 AUTHORS First,M.Kent, Agoulnik,A.I. and Muallem,A.
 TITLE Male infertility y-deletion detection battery
 JOURNAL Patent: US 5776682-A 4 07-JUL-1998;
 FEATURES Location/Qualifiers
 source 1..20
 BASE COUNT 4 a 6 c 3 g 7 t
 ORIGIN
 Query Match 61.0%; Score 12.2; DB 9; Length 20;
 Best Local Similarity 82.4%; Pred. No. 1.3e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 ttcaaccagatggtcat 18
 || ||||| |||||
 Db 1 TCCATCCAGCTGGTCAT 17
 || ||||| |||||
 RESULT 10
 LOCUS AR019114 20 bp DNA PAT 05-DEC-1998
 DEFINITION Sequence 4 from patent US 5783390.
 ACCESSION AR019114
 VERSION AR019114.1 GI:3974228
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 20)
 AUTHORS Svendsen,A., Patkar,S.Anant, Gormsen,E., Okkels,J.Sigurd and Thellersen,M.
 TITLE Lipase variants
 JOURNAL Patent: US 5869438-A 38 09-FEB-1999;
 FEATURES Location/Qualifiers
 source 1..33
 BASE COUNT 10 a 7 c 3 g 13 t
 ORIGIN

AUTHORS First,M.Kent and Agoulnik,A.I.
 TITLE Male infertility y-deletion detection battery
 JOURNAL Patent: US 5783390-A 4 21-JUL-1998;
 FEATURES Location/Qualifiers
 source 1..20
 BASE COUNT 4 a 6 c 3 g 7 t
 ORIGIN
 Query Match 61.0%; Score 12.2; DB 9; Length 20;
 Best Local Similarity 82.4%; Pred. No. 1.3e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 ttcaaccagatggtcat 18
 || ||||| |||||
 Db 1 TCCATCCAGCTGGTCAT 17
 || ||||| |||||
 RESULT 11
 LOCUS AR090906 28 bp DNA PAT 07-SEP-2000
 DEFINITION Sequence 1026 from patent US 5994076.
 ACCESSION AR090906
 VERSION AR090906.1 GI:10017661
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 28)
 AUTHORS Chenchik,A., Johhadze,G. and Bibilashvili,R.
 TITLE Methods of assaying differential expression
 JOURNAL Patent: US 5994076-A 1026 30-NOV-1999;
 FEATURES Location/Qualifiers
 source 1..28
 BASE COUNT 6 a 8 c 5 g 9 t
 ORIGIN
 Query Match 61.0%; Score 12.2; DB 9; Length 28;
 Best Local Similarity 82.4%; Pred. No. 1.3e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 ttcaaccagatggtcat 18
 || ||||| |||||
 Db 3 TTCTACCAGGTGTCAT 19
 || ||||| |||||
 RESULT 12
 LOCUS AR034456 33 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 38 from patent US 5869438.
 ACCESSION AR034456
 VERSION AR034456.1 GI:5950061
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 33)
 AUTHORS Svendsen,A., Patkar,S.Anant, Gormsen,E., Okkels,J.Sigurd and Thellersen,M.
 TITLE Lipase variants
 JOURNAL Patent: US 5869438-A 38 09-FEB-1999;
 FEATURES Location/Qualifiers
 source 1..33
 BASE COUNT 10 a 7 c 3 g 13 t
 ORIGIN
 Query Match 61.0%; Score 12.2; DB 9; Length 33;
 Best Local Similarity 82.4%; Pred. No. 1.3e+05;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttccaaccagatggtca 17
|||||
Db 10 TTTCAACCAAGAGTTAA 26

RESULT 13

AR070401 LOCUS AR070401 33 bp DNA PAT 18-FEB-2000

DEFINITION Sequence 38 from patent US 5892013.

ACCESSION AR070401

VERSION AR070401.1 GI:7221289

KEYWORDS

SOURCE Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 33)

AUTHORS Svendsen,A., Patkar,S.,Anant, Gormsen,E., Clausen,I.,Groth,

TITLE Okkels,J.,Sigurd and Thellersen,M.

JOURNAL Lipase variants

PATENT: US 5892013-A 38 06-APR-1999;

FEATURES

source Location/Qualifiers

1..33

BASE COUNT 10 a 7 c 3 g 13 t

ORIGIN

Query Match 61.08; Score 12.2; DB 9; Length 33;
Best Local Similarity 82.4%; Pred. No. 1.3e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttccaaccagatggtca 17
|||||
Db 10 TTTCAACCAAGAGTTAA 26

RESULT 14

AR026654/c

LOCUS AR026654 34 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 9 from patent US 5856126.

ACCESSION AR026654

VERSION AR026654.1 GI:5937494

KEYWORDS

SOURCE Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 34)

AUTHORS Fukuchi,N., Yamamoto,H., Nagano,M., Kito,M., Tanaka,A., Ishii,K.,

TITLE Kobayashi,T. and Yoshimoto,R.

Peptide having anti-thrombus activity and method of producing the

same

JOURNAL Patent: US 5856126-A 9 05-JAN-1999;

FEATURES

source Location/Qualifiers

1..34

BASE COUNT 5 a 7 c 12 g 10 t

ORIGIN

Query Match 61.08; Score 12.2; DB 9; Length 34;
Best Local Similarity 82.4%; Pred. No. 1.3e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 caaccagatggtcattc 20
|||||
Db 34 CAACCGGAGGACATTC 18

RESULT 15

AX067989/c

LOCUS AX067989 29 bp DNA PAT 19-JAN-2001

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:07 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-18

Perfect score: 20
Sequence: 1 tttaaccagatggtcattc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_est4.*
- 5: gb_est5.*
- 6: gb_est6.*
- 7: gb_est7.*
- 8: gb_est8.*
- 9: gb_est9.*
- 10: gb_est10.*
- 11: gb_est11.*
- 12: gb_est12.*
- 13: gb_est13.*
- 14: gb_est14.*
- 15: gb_est15.*
- 16: gb_est16.*
- 17: gb_est17.*
- 18: gb_est18.*
- 19: gb_est19.*
- 20: gb_est20.*
- 21: gb_est21.*
- 22: gb_est22.*
- 23: gb_est23.*
- 24: gb_est24.*
- 25: gb_est33.*
- 26: gb_est34.*
- 27: gb_est35.*
- 28: gb_est36.*
- 29: gb_est37.*
- 30: gb_est38.*
- 31: gb_est39.*
- 32: gb_est40.*
- 33: em_estba.*
- 34: em_estfun.*
- 35: em_esthum1.*
- 36: em_esthum2.*
- 37: em_esthum3.*
- 38: em_esthum4.*
- 39: em_esthum5.*
- 40: em_esthum6.*
- 41: em_esthum7.*
- 42: em_esthum8.*
- 43: em_esthum9.*

- 44: em_esthum10.*
- 45: em_esthum11.*
- 46: em_esthum12.*
- 47: em_esthum13.*
- 48: em_esthum14.*
- 49: em_esthum15.*
- 50: em_esthum16.*
- 51: em_esthum17.*
- 52: em_esthum18.*
- 53: em_esthum19.*
- 54: em_esthum20.*
- 55: em_esthum21.*
- 56: em_esthum22.*
- 57: em_esthum23.*
- 58: em_esthum24.*
- 59: em_esthum25.*
- 60: em_esthum26.*
- 61: em_esthum27.*
- 62: em_esthum28.*
- 63: em_estin1.*
- 64: em_estin2.*
- 65: em_estin3.*
- 66: em_estin4.*
- 67: em_estin5.*
- 68: em_estom1.*
- 69: em_estom2.*
- 70: em_estov1.*
- 71: em_estov2.*
- 72: em_estpl1.*
- 73: em_estpl2.*
- 74: em_estpl3.*
- 75: em_estpl4.*
- 76: em_estpl5.*
- 77: em_estpl6.*
- 78: em_estpl7.*
- 79: em_estpl8.*
- 80: em_estpl9.*
- 81: em_estpl10.*
- 82: em_estro1.*
- 83: em_estro2.*
- 84: em_estro3.*
- 85: em_estro4.*
- 86: em_estro5.*
- 87: em_estro6.*
- 88: em_estro7.*
- 89: em_estro8.*
- 90: em_estro9.*
- 91: em_estro10.*
- 92: em_estro11.*
- 93: em_estro12.*
- 94: em_estro13.*
- 95: em_estro14.*
- 96: em_estro15.*
- 97: em_estro16.*
- 98: em_estro17.*
- 99: em_estro18.*
- 100: em_estro19.*
- 101: em_estro20.*
- 102: gb_est25.*
- 103: gb_est26.*
- 104: gb_est27.*
- 105: gb_est28.*
- 106: gb_est29.*
- 107: gb_est30.*
- 108: gb_est31.*
- 109: gb_est32.*
- 110: gb_est41.*
- 111: gb_est42.*
- 112: gb_est43.*
- 113: gb_est44.*
- 114: gb_est45.*
- 115: gb_est46.*
- 116: gb_est47.*

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*
 183: gb_est114:*
 184: gb_est115:*
 185: gb_est116:*
 186: gb_est117:*
 187: gb_est118:*
 188: gb_est119:*
 189: gb_est120:*

190: gb_est121:*
 191: gb_est122:*
 192: gb_est123:*
 193: gb_est124:*
 194: gb_est125:*
 195: gb_est126:*
 196: gb_est127:*
 197: gb_est128:*
 198: gb_est129:*
 199: gb_est130:*
 200: gb_est131:*
 201: gb_est132:*
 202: gb_est133:*
 203: gb_est134:*
 204: gb_est135:*
 205: gb_est136:*
 206: gb_est137:*
 207: gb_est138:*
 208: gb_est139:*
 209: gb_est140:*
 210: gb_est141:*
 211: gb_est142:*
 212: gb_est143:*
 213: gb_est144:*
 214: gb_est145:*
 215: gb_est146:*
 216: gb_est147:*
 217: gb_est148:*
 218: gb_est149:*
 219: gb_est150:*
 220: gb_est151:*
 221: gb_est152:*
 222: gb_est153:*
 223: gb_est154:*
 224: gb_est155:*
 225: gb_est156:*
 226: gb_est157:*
 227: gb_est158:*
 228: gb_est159:*
 229: gb_est160:*
 230: gb_est161:*
 231: gb_est162:*
 232: gb_est163:*
 233: gb_est164:*
 234: gb_est165:*
 235: gb_est166:*
 236: gb_est167:*
 237: gb_est168:*
 238: gb_est169:*
 239: gb_est170:*
 240: gb_est171:*
 241: gb_est172:*
 242: gb_est173:*
 243: gb_est174:*
 244: gb_est175:*
 245: gb_est176:*
 246: gb_est177:*
 247: gb_est178:*
 248: gb_est179:*
 249: gb_est180:*
 250: gb_est181:*
 251: gb_est182:*
 252: gb_est183:*
 253: gb_est184:*
 254: gb_est185:*
 255: gb_est186:*
 256: gb_est187:*
 257: gb_est188:*
 258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 1015 Std Error: 0.00
Seq primer: -40UP from Gibco

High quality sequence stop: 1.
Location/Qualifiers

FEATURES
source

1..25
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1836621"
/sex="male"
/sex="male"
/lab_host="DH10B"
/note="vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc. and primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGCGCCCAATTTTTTTTTTTT 3']
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 5 a 8 c 4 g 8 t
ORIGIN

Query Match 62.0%; Score 12.4; DB 17; Length 25;
Best Local Similarity 92.9%; Pred. No. 4.4e+04;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 caaccagatggtca 17
||| ||||| |||||

Db 20 CAATCAGATGGTCA 7

RESULT 2

AZ429959 39 bp DNA GSS 03-OCT-2000
LOCUS
DEFINITION
clone UUGC1M0214N08 F, DNA sequence.

ACCESSION
AZ429959.1 GI:10553972

KEYWORDS
GSS.

SOURCE
house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 39)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)

CONTACT: Robert B. Weiss
University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	12.4	62.0	25	17	AI208968	q929f11.x
c 2	12.4	62.0	39	243	AZ429959	IM0214N08
c 3	12.2	61.0	50	107	AU104028	AU104028
c 4	12.2	61.0	50	107	AU104139	AU104139
c 5	12	60.0	43	242	AZ388563	IM0148F08
c 6	12	60.0	49	16	AI125432	q93d12.x
c 7	12	60.0	49	241	AZ311307	IM0026K11
c 8	11.6	58.0	31	244	AZ453622	IM0255E08
c 9	11.6	58.0	34	247	AZ663664	IM0543J20
c 10	11.6	58.0	35	247	AZ631220	IM0485J13
c 11	11.6	58.0	36	4	AA266819	mz91a02.r
c 12	11.6	58.0	43	2	AA075907	zm74e08.r
c 13	11.6	58.0	46	249	AZ777948	2M0012P13
c 14	11.2	56.0	34	16	AI096045	SMOVL3CAN
c 15	11.2	56.0	42	13	AA936922	om55c08.s
c 16	11.2	56.0	42	249	AZ799790	2M0057M22
c 17	11.2	56.0	43	244	AZ475536	IM0293P23
c 18	11.2	56.0	44	11	AA754427	aj13e09.s
c 19	11.2	56.0	49	10	AA700915	zj44c07.s
c 20	11.2	56.0	50	107	AU104136	AU104136
c 21	11.2	56.0	50	107	AU104138	AU104138
c 22	11	55.0	24	241	AZ317925	IM0036G20
c 23	11	55.0	25	244	AZ449662	IM0248D06
c 24	11	55.0	28	243	AZ427748	IM0209M22
c 25	11	55.0	35	154	BG504713	602551919
c 26	11	55.0	35	250	AZ821491	2M0094H18
c 27	11	55.0	45	189	T80630	yd92c01.s1
c 28	11	55.0	49	2	AA075329	zm82b10.s
c 29	11	55.0	49	107	AU053904	AU053904
c 30	11	55.0	50	31	AV535125	AV535125
c 31	10.8	54.0	29	243	AZ396455	IM0161F06
c 32	10.8	54.0	34	244	AZ451127	IM0250C16
c 33	10.8	54.0	40	14	AA990517	uac3906.s
c 34	10.8	54.0	42	242	AZ388234	IM0148E12
c 35	10.8	54.0	43	243	AZ428834	IM0212F07
c 36	10.8	54.0	44	158	H55075	CHR220014.C
c 37	10.8	54.0	50	114	AW285566	LGI_245.D
c 38	10.6	53.0	26	244	AZ463718	IM0272D16
c 39	10.6	53.0	26	249	AZ780163	2M0017D18
c 40	10.6	53.0	26	249	AZ783434	2M0025G12
c 41	10.6	53.0	30	241	AZ320504	IM0040P03
c 42	10.6	53.0	34	11	AA734076	vs19b04.r
c 43	10.6	53.0	34	13	AA929647	vy91f07.r
c 44	10.6	53.0	34	249	AZ792823	2M0045L06
c 45	10.6	53.0	35	258	TA262E04P	AL484164.T.brucei

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
q929f11.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1836621
3' similar to SW:TEA_LYTPI P02553 TUBULIN ALPHA CHAIN ;, mRNA
sequence.
ACCESSION
AI208968
VERSION
AI208968.1 GI:3770910
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 25)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0214 row: N column: 08
 Seq primer: CGTGTGTAACACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 39.

FEATURES

Location/Qualifiers
 1..39
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0214N08"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: pWD42nv: Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 8 c 15 g 12 t
 ORIGIN

Query Match 62.0%; Score 12.4; DB 243; Length 39;
 Best Local Similarity 92.9%; Pred. No. 4.7e+04;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 tttaaccagatgg 14
 ||| |||||
 Db 20 TTGTAACCATGG 7

RESULT 3

AU104028 50 bp mRNA EST 05-APR-2001
 LOCUS AU104028 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 DEFINITION ADSE00222, mRNA sequence.
 ACCESSION AU104028
 VERSION AU104028.1 GI:13553549
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki.Y., Tsunoda.T., Taira.H., Mizushima-Sugano.J., Sese.J., Hata
 ,K., Suyama.A. and Sugano.S.
 H., Ota.T., Isogai.T., Tanaka.T., Nakamura.Y., Morishita.S., Okubo

TITLE Fine Structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL Unpublished (2001)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp
 Suzuki.Y., Yoshitomo-Nakagawa.K., Maruyama.K., Suyama.A. and Sugano
 ,S. Construction and characterization of a full length-enriched and

FEATURES a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
 Location/Qualifiers
 1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ADSE00222"
 /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 8 a 10 c 16 g 16 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 107; Length 50;
 Best Local Similarity 82.4%; Pred. No. 6.2e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 tttaaccagatgg 17
 ||| |||||
 Db 13 TTTCATCCGTGGTCA 29

RESULT 4

AU104139 50 bp mRNA EST 05-APR-2001
 LOCUS AU104139 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 DEFINITION HEP21348, mRNA sequence.
 ACCESSION AU104139
 VERSION AU104139.1 GI:13553660
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki.Y., Tsunoda.T., Taira.H., Mizushima-Sugano.J., Sese.J., Hata
 ,H., Ota.T., Isogai.T., Tanaka.T., Nakamura.Y., Morishita.S., Okubo
 ,K., Suyama.A. and Sugano.S.
 Fine Structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL Unpublished (2001)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki.Y., Yoshitomo-Nakagawa.K., Maruyama.K., Suyama.A. and Sugano
 ,S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

Location/Qualifiers
 1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEP21348"
 /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 10 a 10 c 15 g 15 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 107; Length 50;
 Best Local Similarity 82.4%; Pred. No. 6.2e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 tttaaccagatgg 18
 ||| |||||
 Db 9 TTCAACCATGGTGTAT 25

RESULT 5

AZ388563/c 43 bp DNA GSS 02-OCT-2000
 LOCUS AZ388563
 DEFINITION 1M0148F08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0148F08 R, DNA sequence.
 ACCESSION AZ388563
 VERSION AZ388563.1 GI:10502271

KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 43)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0148 row: F column: 08
Seq primer: CACACAGGAACACGATGACC
Class: plasmid ends
High quality sequence stop: 43.

FEATURES

source 1..43
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0148F08"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gll4732114/gb1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

16 a 6 c 9 g 12 t

ORIGIN

Query Match 60.0%; Score 12; DB 242; Length 43;
Best Local Similarity 75.0%; Pred. No. 7.7e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tttaacacagatggtcattc 20
||||| ||||| ||||| |||||
Db 39 TTTGAATCACATGTTTATTC 20

RESULT 6

LOCUS A1125432/c 49 bp mRNA EST 11-SEP-1998
DEFINITION q93d12.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1737047 3' similar to SW:LI11_NYCCO P08548 LINE-1 REVERSE TRANSCRIPTASE HOMOLOG. ;contains L1.t3 L1 repetitive element ;, mRNA sequence.

ACCESSION A1125432
VERSION A1125432.1 GI:3593946
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 49)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES

source 1..49
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1737047"
/clone.lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - Oligo(dT) primer [5', TGTTACCAATCTGAAGTGGGCGCGCCAAATTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

8 a 8 c 14 g 19 t

Query Match 60.0%; Score 12; DB 16; Length 49;
Best Local Similarity 100.0%; Pred. No. 7.8e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 tcaaccagatgg 14
||||| ||||| |||||
Db 20 TCAACCAAGATGG 9

RESULT 7

LOCUS AZ311307 49 bp DNA GSS 29-SEP-2000
DEFINITION IM0026K11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0026K11 R, DNA sequence.

ACCESSION AZ311307
VERSION AZ311307.1 GI:10354140
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 49)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0026 row: K column: 11
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 49.

FEATURES

Location/Qualifiers
 1. .49
 source

/organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0026K11"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 8 a 10 c 8 g 23 t
 ORIGIN

Query Match 60.0%; Score 12; DB 241; Length 49;
 Best Local Similarity 75.0%; Pred. No. 7.8e+04;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ttccaaccagatggcttc 20
 ||||| ||||| |||||
 Db 12 TTTCACCCAGATTCTATTC 31

RESULT 8

AZ453622/c
 LOCUS AZ453622 31 bp DNA GSS 04-OCT-2000
 DEFINITION IM0255E08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0255E08 F, DNA sequence.
 ACCESSION AZ453622
 VERSION AZ453622.1 GI:10611604
 KEYWORDS GSS
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 31)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D.,Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0255 row: E column: 08
 Seq primer: CGTTCTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 31.

FEATURES

Location/Qualifiers
 1. .31
 source
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0255E08"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 9 a 8 c 7 g 7 t
 ORIGIN

Query Match 58.0%; Score 11.6; DB 244; Length 31;
 Best Local Similarity 77.8%; Pred. No. 1.2e+05;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttccaaccagatggctcat 18
 ||||| ||||| |||||
 Db 18 TTTCACCCAGATGGGCAT 1

RESULT 9

AZ663664/c
 LOCUS AZ663664 34 bp DNA GSS 14-DEC-2000
 DEFINITION IM0543J20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0543J20 F, DNA sequence.
 ACCESSION AZ663664
 VERSION AZ663664.1 GI:11800810
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 34)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0543 row: J column: 20
 Seq primer: CGTTGTAAACGACGCCACGT
 Class: plasmid ends
 High quality sequence stop: 34.

FEATURES source

```

1. .34
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUGC1M0543J20"
  /clone_lib="Mouse 10kb plasmid UUGC1M library"
  /sex="Male"
  /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
  /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      8 a      7 c      7 g      12 t
ORIGIN

```

Query Match 58.0%; Score 11.6; DB 247; Length 34;
 Best Local Similarity 77.8%; Pred. No. 1.2e+05;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 3 tcaaccagatggtcattc 20
   | ||||| ||||| |||||
Db 23 TGAACCAATGGCGATTC 6

```

RESULT 10 AZ6311220

LOCUS AZ6311220 35 bp DNA GSS 13-DEC-2000
 DEFINITION 1M0485J13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0485J13 F, DNA sequence.
 ACCESSION AZ6311220
 VERSION AZ6311220.1 GI:11753410
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 35)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0485 row: J column: 13
 Seq primer: CGTTGTAAACGACGCCACGT
 Class: plasmid ends
 High quality sequence stop: 35.

FEATURES source

```

1. .35
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUGC1M0485J13"
  /clone_lib="Mouse 10kb plasmid UUGC1M library"
  /sex="Male"
  /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
  /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      14 a      6 c      5 g      10 t
ORIGIN

```

Query Match 58.0%; Score 11.6; DB 247; Length 35;
 Best Local Similarity 77.8%; Pred. No. 1.2e+05;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 ttcaaccagatggtcat 18
   | ||||| ||||| |||
Db 15 TTTCAGGAGATGATCCT 32

```

RESULT 11 AA266819/c

LOCUS AA266819 36 bp mRNA EST 21-MAR-1997
 DEFINITION mz91a02.r1 Soares mouse lymph node NbMLN Mus musculus cDNA clone IMAGE:720746 5' similar to gb:DI0061 Mouse mRNA for DNA topoisomerase I (MOUSE);, mRNA sequence.
 ACCESSION AA266819
 VERSION AA266819.1 GI:1903574
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0012 row: P column: 13
 Seq primer: CACACAGGAACAGCATGACC
 Class: plasmid ends
 High quality sequence stop: 46.

FEATURES

source

1. .46
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0012P13"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydronamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g14732114|g9|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 ORIGIN

20 a 8 c 11 g 7 t

Query Match 58.0%; Score 11.6; DB 249; Length 46;
 Best Local Similarity .77.8%; Pred. No. 1.2e+05;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tttaaccagatggtcat 18
 ||||| ||| ||| |||
 Db 34 TTTCATCAGCTGCACAT 17

RESULT 14
 AI096045/c

LOCUS SW02L3CAN20F05 Onchocerca volvulus infective larva CDNA
 DEFINITION (SAW94WL-OVL3) Onchocerca volvulus cdna clone onch1306 5' similar to TR:Q25606 Q25606 LARVAL 18 KDA PROTEIN.; mRNA sequence.

ACCESSION AI096045

VERSION AI096045

KEYWORDS EST.

SOURCE Onchocerca volvulus.

ORGANISM Onchocerca volvulus.

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Onchocerca.

AUTHORS 1 (bases 1 to 34)

Williams, S.A., Lizotte-Waniewski, M., Laney, S., Wenhong, L., Hillier, L., Allen, M., Bowles, L., Geisel, S., Jost, S., Kucaba, T., Martin, J., Steptoe, M., Theising, B., White, F., Wyllie, T., Chappell, J., Person, B., Gibbons, M., Harvey, N., Pape, D., Chamberlain, A., Morales, R., Schurk, R., Ritter, E., Kohn, S., Underwood, K. and Marra, M.

TITLE Molecular Parasitology OVL3

JOURNAL Unpublished (1998)

COMMENT Contact: Steven A. Williams

Molecular Parasitology

Smith College Department of Biological Sciences

Department of Biological Sciences, Clark Science Center, Smith College, Northampton, MA, 01063, USA
 Tel: 4135853826
 Fax: 4135853786
 Email: genome@smith.edu

The library was constructed by Wenhong Lu. The library is available from Dr. S.A. Williams, email genome@smith.edu When requesting this clone from Dr. Williams, please reference the Williams lab clone id - SW02L3CAN20F05

Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: T3 ET from Amersham
 High quality sequence stop: 1.

FEATURES

source

1. .34
 /organism="Onchocerca volvulus"
 /strain="Sierra Leone"
 /db_xref="taxon:6282"
 /clone="onch1306"
 /clone_lib="Onchocerca volvulus infective larva CDNA (SAW94WL-OVL3)"
 /lab_host="XL1-Blue MRP"
 /note="Vector: lambda UniZap XR; Site_1: EcoR I; Site_2: Xho I; Cutaneous filarial nematode parasite of humans. mRNA was prepared from third stage infective larvae of Onchocerca volvulus isolated from mosquitoes 10 days after infection and converted to double stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNapol I. The library had 1.8 x 10⁵ independent recombinants and average insert size was 900 base pairs. The library was constructed by Wenhong Lu. The library is available from Dr. S.A. Williams, email genome@smith.edu."

BASE COUNT 12 a 6 c 6 g 10 t

ORIGIN

Query Match 56.0%; Score 11.2; DB 16; Length 34;
 Best Local Similarity 81.2%; Pred. No. 1.9e+05;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 caaccagatggtcatt 19
 ||| || ||||| ||
 Db 27 CAATCAATGTCGTT 12

RESULT 15

AA936922

LOCUS

DEFINITION

AA936922 42 bp mRNA

EST

AA936922.1 GI:3094956

KEYWORDS

SOURCE

ORGANISM

Human.

Human sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 42)

NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapb-f@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

CDNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CCGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 753 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES
source

Location/Qualifiers
1. .42
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1551086"
/clone_lib="NCI_CGAP_GC4"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(UT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 14 a 7 c 11 g 10 t
ORIGIN

Query Match 56.0%; Score 11.2; DB 13; Length 42;
Best Local Similarity 81.2%; Pred. No. 1.9e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 tcaaccagatggctcat 18
||| | |||||
Db 22 TCATCATGATGGTCAT 37

Search completed: October 2, 2001, 15:01:08
Job time: 10837 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 12:00:36 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-3

Perfect score: 20

Sequence: 1 ccgcgggtcacagtgtcg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:**

1: gb_bal:**

2: gb_ba2:**

3: gb_ba3:**

4: gb_in1:**

5: gb_in2:**

6: gb_in3:**

7: gb_om:**

8: gb_ov:**

9: gb_pat1:**

10: gb_pat2:**

11: gb_ph:**

12: gb_p11:**

13: gb_p12:**

14: gb_p13:**

15: gb_p14:**

16: gb_bal:**

17: gb_ba2:**

18: em_fun:**

19: em_htgo_hum:**

20: em_htgo_inv:**

21: em_htgo_rod:**

22: em_htg_hum1:**

23: em_htg_hum2:**

24: em_htg_hum3:**

25: em_htg_hum4:**

26: em_htg_hum5:**

27: em_htg_hum6:**

28: em_htg_hum7:**

29: em_htg_hum8:**

30: em_htg_inv1:**

31: em_htg_inv2:**

32: em_htg_other:**

33: em_htg_rod:**

34: em_hum1:**

35: em_hum2:**

36: em_hum3:**

37: em_hum4:**

38: em_hum5:**

39: em_hum6:**

40: em_hum7:**

41: em_in:**

42: em_om:**

43: em_or:**

44: em_ov:**
45: em_pat:**
46: em_ph:**
47: em_pl:**
48: em_ro:**
49: em_sts:**
50: em_sy:**
51: em_un:**
52: em_vi:**
53: gb_sts1:**
54: gb_sts2:**
55: gb_sts3:**
56: gb_sy:**
57: gb_un:**
58: gb_vil:**
59: gb_v12:**
60: gb_htg1:**
61: gb_htg2:**
62: gb_htg3:**
63: gb_htg4:**
64: gb_htg5:**
65: gb_htg6:**
66: gb_htg7:**
67: gb_htg8:**
68: gb_htg9:**
69: gb_htg10:**
70: gb_htg11:**
71: gb_htg12:**
72: gb_htg13:**
73: gb_htg14:**
74: gb_htg15:**
75: gb_htg16:**
76: gb_htg17:**
77: gb_htg18:**
78: gb_htg19:**
79: gb_htg20:**
80: gb_htg21:**
81: gb_htg22:**
82: gb_htg23:**
83: gb_htg24:**
84: gb_htg25:**
85: gb_pr1:**
86: gb_pr2:**
87: gb_pr3:**
88: gb_pr4:**
89: gb_pr5:**
90: gb_pr6:**
91: gb_pr7:**
92: gb_pr8:**
93: gb_pr9:**
94: gb_rol:**
95: gb_ro2:**
96: gb_in4:**
97: gb_pr10:**
98: em_ba3:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	13.2	66.0	33	10	I92389
2	13.2	66.0	50	9	AX043850
3	12.6	63.0	37	10	AX090386
4	12.6	63.0	41	10	AX090365
5	12.4	62.0	45	94	MWTCRAIC7
6	12.2	61.0	29	9	AX026028
7	12.2	61.0	30	10	E28465
8	12.2	61.0	40	9	AR095613
					I92389 Sequence 41
					AX043850 Sequence
					AX090386 Sequence
					AX090365 Sequence
					X70713 M.musculus
					AX026028 Sequence
					E28465 Hyaluronate
					AR095613 Sequence

9 12.2 61.0 40 9 AR095615
 10 12.2 61.0 41 9 AX043844
 11 12.2 61.0 44 10 AX090361
 12 12.2 61.0 45 9 AX043846
 13 11.8 59.0 16 9 AR041138
 14 11.8 59.0 22 10 I43055
 15 11.8 59.0 29 9 AR041128
 16 11.8 59.0 30 9 AR041129
 17 11.8 59.0 30 9 AR041130
 18 11.8 59.0 35 10 I31649
 19 11.8 59.0 36 10 I21472
 20 11.8 59.0 44 10 I43051
 21 11.6 58.0 24 9 AR017555
 22 11.6 58.0 24 10 I21733
 23 11.6 58.0 25 9 A58772
 24 11.6 58.0 25 10 I19529
 25 11.6 58.0 25 10 I19530
 26 11.6 58.0 25 10 I19532
 27 11.6 58.0 25 10 I86209
 28 11.6 58.0 25 10 I86210
 29 11.6 58.0 25 10 I86212
 30 11.6 58.0 26 9 A58770
 31 11.6 58.0 26 9 AR017553
 32 11.6 58.0 26 10 I19534
 33 11.6 58.0 26 10 I21731
 34 11.6 58.0 26 10 I86214
 35 11.6 58.0 30 9 AR049374
 36 11.6 58.0 30 9 AR095535
 37 11.6 58.0 30 9 AX033429
 38 11.6 58.0 50 9 A58773
 39 11.4 57.0 18 9 AR078873
 40 11.4 57.0 22 10 I76917
 41 11.4 57.0 25 9 A65881
 42 11.4 57.0 25 9 AX042765
 43 11.4 57.0 37 9 S45430
 44 11.2 56.0 20 9 AR067323
 45 11.2 56.0 20 9 AR073339

ALIGNMENTS

RESULT 1
 192389
 LOCUS I92389 33 bp DNA PAT 01-DEC-1998
 DEFINITION Sequence 41 from patent US 5728519.
 ACCESSION I92389
 VERSION I92389.1 GI:3936859
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 33)
 AUTHORS Levenbook I.S., Chumakov, K.M., Norwood, L.P. and Roninson, I.
 TITLE Assay for virulent revertants of attenuated live vaccines and kits therefor
 JOURNAL Patent: US 5728519-A 41 17-MAR-1998;
 FEATURES Location/Qualifiers
 source 1..33
 BASE COUNT 6 a 7 c 9 g 11 t
 ORIGIN
 Query Match 66.0%; Score 13.2; DB 10; Length 33;
 Best Local Similarity 83.3%; Pred. No. 5.2e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ccgcgggttcacagtggc 18
 ||||| ||| ||| ||
 Db 9 CCGCGGGTTCAAAGTAGT 26

RESULT 2
 AX043850
 LOCUS AX043850 50 bp DNA PAT 23-NOV-2000
 DEFINITION Sequence 57 from Patent WO0063391.
 ACCESSION AX043850
 VERSION AX043850.1 GI:11342435
 KEYWORDS
 ORGANISM synthetic construct.
 SOURCE synthetic construct.
 REFERENCE 1 (bases 1 to 50)
 AUTHORS Savidge, B., Lassner, M.W., Weiss, J.D. and Post-Beitemmiller, D.
 TITLE Nucleic acid sequences to proteins involved in tocopherol synthesis
 JOURNAL Patent: WO 0063391-A 57 26-OCT-2000;
 FEATURES Location/Qualifiers
 source 1..50
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Oligonucleotide"
 BASE COUNT 9 a 16 c 14 g 11 t
 ORIGIN
 Query Match 66.0%; Score 13.2; DB 9; Length 50;
 Best Local Similarity 83.3%; Pred. No. 4.7e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ccgcgggttcacagtggc 18
 ||||| ||| ||| |||
 Db 5 CCGCGGCCGCACATGGT 22

RESULT 3
 AX090386
 LOCUS AX090386 37 bp DNA PAT 21-MAR-2001
 DEFINITION Sequence 79 from Patent WO0116308.
 ACCESSION AX090386
 VERSION AX090386.1 GI:13444245
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 37)
 AUTHORS Lassner, M. and van Eenennaam, A.
 TITLE Plant sterol acyltransferases
 JOURNAL Patent: WO 0116308-A 79 08-MAR-2001;
 FEATURES Location/Qualifiers
 source 1..37
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic oligonucleotide primer"
 BASE COUNT 8 a 12 c 11 g 6 t
 ORIGIN
 Query Match 63.0%; Score 12.6; DB 10; Length 37;
 Best Local Similarity 78.9%; Pred. No. 1e+05;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ccgcgggttcacagtggc 19
 ||||| ||| ||| |||
 Db 5 CCGCGGCCGCACATGGC 23

RESULT 4
 AX090365
 LOCUS AX090365 41 bp DNA PAT 21-MAR-2001
 DEFINITION Sequence 58 from Patent WO0116308.
 ACCESSION AX090365
 VERSION AX090365.1 GI:13444226
 KEYWORDS

SOURCE
ORGANISM synthetic construct.
artificial sequence.

REFERENCE
1 (bases 1 to 41)
Lassner,M. and van Eenennaam,A.
TITLE Plant sterol acyltransferases
JOURNAL Patent: WO 0116308-A 58 08-MAR-2001;
MONSANTO COMPANY (US)

FEATURES
source Location/Qualifiers
1..41
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic oligonucleotide primer"
8 t

BASE COUNT 5 a 13 c 15 g 8 t

ORIGIN

Query Match 63.0%; Score 12.6; DB 10; Length 41;
Best Local Similarity 78.9%; Pred. No. 9.9e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccgcgggctcacagtgtc 19
||||| | |||| | |||

Db 5 CCgcggccgcacacatggc 23

RESULT 5
MMTCRAIC7/c

LOCUS MMTCAIC7 45 bp mRNA ROD 15-MAR-1993

DEFINITION M.musculus mRNA for T-cell receptor, V-J alpha junction (A3/IIC7).

ACCESSION X70713

VERSION X70713.1 GI:288017

KEYWORDS J-alpha gene segment; junction; T-cell receptor; V-alpha gene segment.

SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Casanova,J.

REFERENCE
1 (bases 1 to 45)
Direct Submission
Submitted (15-JAN-1993) J. Casanova, INSERM U132, Developpement
Normal et Pathologique du Systeme Immunitaire, Necker, Pavillon
Kirmisson, 149 rue de Sevres, 75743 Paris Cedex, FRANCE

REFERENCE
2 (bases 1 to 45)
Casanova,J.L., Martinon,F., Gournier,H., Barra,C., Pannetier,C.,
Regnault,A., Kourilsky,P., Cerottini,J.C. and Maryanski,J.L.
TITLE T cell receptor selection by and recognition of two class I major
histocompatibility complex-restricted antigenic peptides that
differ at a single position
J. Exp. Med. 177 (3), 811-820 (1993)

JOURNAL 93171821

MEDLINE

FEATURES
source Location/Qualifiers
1..45
/organism="Mus musculus"
/strain="DBA2"
/isolate="A3/IIC7"
/db_xref="taxon:10090"
/haplotype="H2-d"
/tissue_type="T cell"
/cell_type="CTL"
<1..>45
/note="T-cell receptor, V-J alpha junction"
12 t

BASE COUNT 6 a 9 c 18 g 12 t

ORIGIN

Query Match 62.0%; Score 12.4; DB 94; Length 45;
Best Local Similarity 92.9%; Pred. No. 1.2e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ccgcgggctcacag 14
||||| |||||||

Db 18 CCgcggcgtcacag 5

RESULT 6
AX026028/c

LOCUS AX026028 29 bp DNA PAT 16-SEP-2000

DEFINITION Sequence 9 from Patent WO0031239.

ACCESSION AX026028

VERSION AX026028.1 GI:10187483

KEYWORDS synthetic construct.

SOURCE synthetic construct.
artificial sequence.

ORGANISM 1 (bases 1 to 29)
Eshhar,Z., Bolhuis,R.L. and Willemssen,R.A.
TITLE Immune cells having predefined biological specificity, comprising
chimeric t cell receptor
JOURNAL Patent: WO 0031239-A 9 02-JUN-2000;
ESHAR ZELIG (IL); YEDA RES & DEV (IL); BOLHUIS REINDER L H (NL)
; WILLEMSSEN RALPH A (NL)

FEATURES
source Location/Qualifiers
1..29
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic DNA"
4 t

BASE COUNT 4 a 10 c 11 g 4 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 29;
Best Local Similarity 82.4%; Pred. No. 1.7e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 cgggctcacagtgtcg 20
| |||| | |||| | |||

Db 24 CCgcgtcacagtgtcg 8

RESULT 7
E28465/c

LOCUS E28465 30 bp DNA PAT 07-FEB-2001

DEFINITION Hyaluronate synthase promoter DNA.

ACCESSION E28465

VERSION E28465.1 GI:13018357

KEYWORDS JP 1999196875-A/13.

SOURCE unclassified.

ORGANISM unclassified.
1 (bases 1 to 30)
Yoichi,Y.N.I.I. and Kimata.
TITLE Hyaluronate synthase promoter DNA
JOURNAL Patent: JP 1999196875-A 13 27-JUL-1999;
SEIKAGAKU KOGYO CO LTD

COMMENT OS Unidentified
PN JP 1999196875-A/13
PD 27-JUL-1999
PF 14-JAN-1998 JP 1998006191
PR YOICHI YAMADA,NAOKI ITANO,KOJI KIMATA
PI C12N15/09,C12N9/00,C12Q1/68//C12N15/09,C12R1:91,C12N15/00,
PC C12N15/00,C12R1:91
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
1..30
FT source /organism="Unidentified".

FEATURES
source Location/Qualifiers
1..30
/organism="unidentified"
/db_xref="taxon:32644"
7 a 11 c 7 g 5 t

BASE COUNT 7 a 11 c 7 g 5 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 10; Length 30;
 Best Local Similarity 82.4%; Pred. No. 1.7e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ccgcgggctcacagtgg 18
 |||| ||||| |||||
 Db 18 CCGGTGCTCATGTGT 2

RESULT 8
 AR095613
 LOCUS AR095613 40 bp DNA PAT 08-SEP-2000
 DEFINITION Sequence 8 from patent US 6004793.
 ACCESSION AR095613
 VERSION AR095613.1 GI:10023640
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 40)
 AUTHORS Lunnen,K.D., Dalton,M.A., Wilson,G.G. and Xu,S.
 TITLE Method for cloning and producing the AwaI restriction endonuclease in E. coli and purification of the recombinant AwaI restriction endonuclease
 JOURNAL Patent: US 6004793-A 8 21-DEC-1999;
 FEATURES Location/Qualifiers
 source 1..40
 /organism="unknown"

BASE COUNT 12 a 12 c 9 g 7 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 40;
 Best Local Similarity 82.4%; Pred. No. 1.6e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccgcgggctcacagtgg 17
 || ||||| |||||
 Db 8 CCGCGGCTCAAGTTG 24

RESULT 9
 AR095615
 LOCUS AR095615 40 bp DNA PAT 08-SEP-2000
 DEFINITION Sequence 10 from patent US 6004793.
 ACCESSION AR095615
 VERSION AR095615.1 GI:10023644
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 40)
 AUTHORS Lunnen,K.D., Dalton,M.A., Wilson,G.G. and Xu,S.
 TITLE Method for cloning and producing the AwaI restriction endonuclease in E. coli and purification of the recombinant AwaI restriction endonuclease
 JOURNAL Patent: US 6004793-A 10 21-DEC-1999;
 FEATURES Location/Qualifiers
 source 1..40
 /organism="unknown"

BASE COUNT 12 a 12 c 9 g 7 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 40;
 Best Local Similarity 82.4%; Pred. No. 1.6e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccgcgggctcacagtgg 17
 || ||||| |||||
 Db 8 CCGCGGCTCAAGTTG 24

RESULT 10
 AX043844
 LOCUS AX043844 41 bp DNA PAT 23-NOV-2000
 DEFINITION Sequence 51 from Patent WO0063391.
 ACCESSION AX043844
 VERSION AX043844.1 GI:11342429
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 41)
 AUTHORS Savidge,B., Lassner,M.W., Weiss,J.D. and Post-Beittenmiller,D.
 TITLE Nucleic acid sequences to proteins involved in tocopherol synthesis
 JOURNAL Patent: WO 0063391-A 51 26-OCT-2000;
 FEATURES Location/Qualifiers
 source 1..41
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Oligonucleotide"

BASE COUNT 6 a 13 c 11 g 11 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 41;
 Best Local Similarity 82.4%; Pred. No. 1.6e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccgcgggctcacagtgg 17
 ||||| ||||| |||||
 Db 5 CCGCGCGGCACATGG 21

RESULT 11
 AX090361
 LOCUS AX090361 44 bp DNA PAT 21-MAR-2001
 DEFINITION Sequence 54 from Patent WO0116308.
 ACCESSION AX090361
 VERSION AX090361.1 GI:13444222
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 44)
 AUTHORS Lassner,M. and van Eenennaam,A.
 TITLE Plant sterol acyltransferases
 JOURNAL Patent: WO 0116308-A 54 08-MAR-2001;
 FEATURES Location/Qualifiers
 source 1..44
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic oligonucleotide primer"

BASE COUNT 13 a 11 c 14 g 6 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 10; Length 44;
 Best Local Similarity 82.4%; Pred. No. 1.6e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccgcgggctcacagtgg 17
 ||||| ||||| |||||
 Db 5 CCGCGCGGCACATGG 21

RESULT 12
 AX043846
 LOCUS AX043846 45 bp DNA PAT 23-NOV-2000
 DEFINITION Sequence 53 from Patent WO0063391.
 ACCESSION AX043846
 VERSION AX043846.1 GI:11342431

JOURNAL FEATURES	Patent: US 5631130-A 38 20-MAY-1997;									
SOURCE	Location/Qualifiers 1. .22 /organism="unknown"									
BASE COUNT	4	a	9	c	4	g	5	t		
ORIGIN										
Query Match	59.0%; Score 11.8; DB 10; Length 22;									
Best Local Similarity	86.7%; Pred. No. 2.9e+05;									
Matches	13;	Conservative	0;	Mismatches	2;	Indels	0;	caps	0;	
Qy	1	ccgcggggctcacagt	15							
Db	7	CCGCACGCTCACAGT	21							

RESOLUTION	LOCUS	DEFINITION	SEQUENCE	29 bp	DNA	PAT	29-SEP-1999
AR041128	AR041128	Sequence 3	from patent	US 5811269.			
AR041128	AR041128	ACCESSION	AR041128				
AR041128	AR041128	VERSION	AR041128.1	GI:5961624			
AR041128	AR041128	KEYWORDS					
AR041128	AR041128	SOURCE		Unknown.			

```

REFERENCE 1 (bases 1 to 29)
AUTHORS Nadeau,J.G., Dean,C.H., Schram,J.L., Howard,D.R., Dey,M.S. and Wright,D.J.
TITLE Detection of mycobacteria by multiplex nucleic acid amplification
JOURNAL Patent: US 5811269-A 3 23-SEP-1998;
FEATURES Location/Qualifiers
source 1..29
/organism="unknown"
BASE COUNT 7 a 9 c 6 t
ORIGIN
Query Match 59.0%; Score 11.8; DB 9; Length 29;
Best Local Similarity 86.7%; Pred. No. 2.7e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ccgcggggtcacagt 15
||| |||||
Db 14 CCGCATGCTCACACT 28

```

Search completed: October 2, 2001, 15:56:35
Job time: 14159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 2, 2001, 12:00:31 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-3
Perfect score: 20
Sequence: 1 ccgcgggtcacagtggtcg 20
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: em_esthum29:*
170: em_esthum30:*
171: em_esthum31:*
172: em_esthum32:*
173: em_esthum33:*
174: em_estom3:*
175: em_estpl11:*
176: em_estro21:*
177: em_estro22:*
178: em_estro23:*
179: em_hic:*
180: gb_est107:*
181: gb_est108:*
182: gb_est109:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_hic:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	14.4	72.0	50	107	AU106328	AU106328 AU106328
2	13.6	68.0	46	3	AA196905	AA196905 zq09b06.r
3	13.2	66.0	50	107	AU107932	AU107932 AU107932
4	12.8	64.0	40	244	AZ453043	AZ453043 AU107932
5	12.8	64.0	41	12	AA822940	AA822940 AU107932
6	12.6	63.0	25	24	AA822940	AA822940 VP30903.r
7	12.6	63.0	25	24	AA822940	AA822940 WH6510.r
8	12.6	63.0	46	244	AA845818	AA845818 ak85f01.s
9	12.6	63.0	46	244	AZ465895	AZ465895 1M0276F12
10	12.6	63.0	50	250	AZ834614	AZ834614 2M0117I21
11	12.6	60.0	44	258	TA233A12Q	TA233A12Q T. brucei
12	12.6	60.0	50	107	AU106948	AU106948 AU106948
13	11.6	58.0	22	258	TA140A04P	TA140A04P T. brucei
14	11.6	58.0	28	24	AI735009	AI735009 as44604.x
15	11.6	58.0	37	7	AA466917	AA466917 vf10904.r
16	11.6	58.0	46	9	AA607492	AA607492 vm61h02.r
17	11.6	58.0	46	16	AI149037	AI149037 qz75409.x
18	11.4	57.0	50	107	AU106912	AU106912 AU106912
19	11.4	57.0	29	249	AZ766277	AZ766277 1M0563A15
20	11.4	57.0	50	107	AU104386	AU104386 AU104386
21	11.2	56.0	25	258	TA205H11Q	TA205H11Q T. brucei
22	11.2	56.0	31	242	AZ345566	AZ345566 1M0080I14
23	11.2	56.0	37	23	AI684291	AI684291 tw97b09.x
24	11.2	56.0	37	250	AZ812781	AZ812781 2M0079B17
25	11.2	56.0	38	175	BG292849	BG292849 602389872
26	11.2	56.0	50	107	AU102352	AU102352 AU102352
27	11.2	56.0	50	107	AU102356	AU102356 AU102356
28	11.2	56.0	50	107	AU104505	AU104505 AU104505
29	11.2	56.0	32	243	AZ404438	AZ404438 1M0173I16
30	11.2	56.0	41	247	AZ647118	AZ647118 1M0513023
31	11.2	55.0	44	166	BES11212	BES11212 601057922
32	11.2	55.0	46	144	BF123084	BF123084 601761846
33	11.2	55.0	46	242	AZ388569	AZ388569 1M0148G09
34	11.2	55.0	50	107	AU103426	AU103426 AU103426
35	11.2	55.0	50	107	AU103606	AU103606 AU103606
36	11.2	55.0	50	107	AU103608	AU103608 AU103608
37	11.2	55.0	50	107	AU103614	AU103614 AU103614
38	11.2	55.0	50	107	AU103617	AU103617 AU103617
39	11.2	55.0	50	107	AU104542	AU104542 AU104542
40	11.2	55.0	50	107	AU107007	AU107007 AU107007
41	10.8	54.0	25	249	AZ770438	AZ770438 1M0572G03
42	10.8	54.0	27	244	AZ455866	AZ455866 1M0258K21
43	10.8	54.0	27	244	AZ463607	AZ463607 1M0272K02
44	10.8	54.0	35	245	AZ497161	AZ497161 1M0333G22
45	10.8	54.0	38	246	AZ591206	AZ591206 1M0401K03

ALIGNMENTS

Result No.	Score	Match	Length	DB	ID	Description
1	14.4	72.0	50	107	AU106328	AU106328 AU106328
2	13.6	68.0	46	3	AA196905	AA196905 zq09b06.r
3	13.2	66.0	50	107	AU107932	AU107932 AU107932
4	12.8	64.0	40	244	AZ453043	AZ453043 AU107932
5	12.8	64.0	41	12	AA822940	AA822940 AU107932
6	12.6	63.0	25	24	AA822940	AA822940 VP30903.r
7	12.6	63.0	25	24	AA822940	AA822940 WH6510.r
8	12.6	63.0	46	244	AA845818	AA845818 ak85f01.s
9	12.6	63.0	46	244	AZ465895	AZ465895 1M0276F12
10	12.6	63.0	50	250	AZ834614	AZ834614 2M0117I21
11	12.6	60.0	44	258	TA233A12Q	TA233A12Q T. brucei
12	12.6	60.0	50	107	AU106948	AU106948 AU106948
13	11.6	58.0	22	258	TA140A04P	TA140A04P T. brucei
14	11.6	58.0	28	24	AI735009	AI735009 as44604.x
15	11.6	58.0	37	7	AA466917	AA466917 vf10904.r
16	11.6	58.0	46	9	AA607492	AA607492 vm61h02.r
17	11.6	58.0	46	16	AI149037	AI149037 qz75409.x
18	11.4	57.0	50	107	AU106912	AU106912 AU106912
19	11.4	57.0	29	249	AZ766277	AZ766277 1M0563A15
20	11.4	57.0	50	107	AU104386	AU104386 AU104386
21	11.2	56.0	25	258	TA205H11Q	TA205H11Q T. brucei
22	11.2	56.0	31	242	AZ345566	AZ345566 1M0080I14
23	11.2	56.0	37	23	AI684291	AI684291 tw97b09.x
24	11.2	56.0	37	250	AZ812781	AZ812781 2M0079B17
25	11.2	56.0	38	175	BG292849	BG292849 602389872
26	11.2	56.0	50	107	AU102352	AU102352 AU102352
27	11.2	56.0	50	107	AU102356	AU102356 AU102356
28	11.2	56.0	50	107	AU104505	AU104505 AU104505
29	11.2	56.0	32	243	AZ404438	AZ404438 1M0173I16
30	11.2	56.0	41	247	AZ647118	AZ647118 1M0513023
31	11.2	55.0	44	166	BES11212	BES11212 601057922
32	11.2	55.0	46	144	BF123084	BF123084 601761846
33	11.2	55.0	46	242	AZ388569	AZ388569 1M0148G09
34	11.2	55.0	50	107	AU103426	AU103426 AU103426
35	11.2	55.0	50	107	AU103606	AU103606 AU103606
36	11.2	55.0	50	107	AU103608	AU103608 AU103608
37	11.2	55.0	50	107	AU103614	AU103614 AU103614
38	11.2	55.0	50	107	AU103617	AU103617 AU103617
39	11.2	55.0	50	107	AU104542	AU104542 AU104542
40	11.2	55.0	50	107	AU107007	AU107007 AU107007
41	10.8	54.0	25	249	AZ770438	AZ770438 1M0572G03
42	10.8	54.0	27	244	AZ455866	AZ455866 1M0258K21
43	10.8	54.0	27	244	AZ463607	AZ463607 1M0272K02
44	10.8	54.0	35	245	AZ497161	AZ497161 1M0333G22
45	10.8	54.0	38	246	AZ591206	AZ591206 1M0401K03

JOURNAL
COMMENT

Unpublished (2001)
Contact: Yuraka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp

S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source

Location/Qualifiers
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT00883"
/clone_lib="Sugano Homo sapiens cDNA library"
6 a 12 c 22 g 10 t

BASE COUNT
ORIGIN

Query Match 72.0% Score 14.4; DB 107; Length 50;
Best Local Similarity 93.8%; Pred. No. 4.8e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ggggtcacagtggctg 20
|||||
Db 22 GGGCGCACAGTGGTCG 37

RESULT 2
LOCUS

AA196905/c
DEFINITION
zq09b06.r1 Stratagene muscle 937209 Homo sapiens cDNA clone
IMAGE:629171 5' similar to FR:E36241 E36241 PMS PROTEIN.; mRNA
sequence.

ACCESSION
VERSION

AA196905
AA196905.1 GI:1792496
EST
KEYWORDS

SOURCE
ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 46)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maria, M., Martin
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wyllie, T., Waterston, R. and Wilson, R.

TITLE
JOURNAL

Unpublished (1997)
Contact: Willson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone; similarity on wrong strand
Insert length: 1100 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..46
/organism="Homo sapiens"
/db_xref="GDB:5049405"
/db_xref="taxon:9606"
/clone="IMAGE:629171"
/clone_lib="Stratagene muscle 937209"
/tissue_type="muscle"
/dev_stage="adult"

FEATURES
source

Location/Qualifiers
1..46
/organism="Homo sapiens"
/db_xref="GDB:5049405"
/db_xref="taxon:9606"
/clone="IMAGE:629171"
/clone_lib="Stratagene muscle 937209"
/tissue_type="muscle"
/dev_stage="adult"
/note="Organ: skeletal muscle; Vector: pBluescript SK-;

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

Unpublished (2001)
Contact: Yuraka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally.
 Primer: Oligo dt. Skeletal muscle from patient with
 malignant hyperthermia. Average insert size: 1.0 kb;
 Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGG
 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT
 ORIGIN

12 a 15 c 14 g 5 t

Query Match 68.0%; Score 13.6; DB 3; Length 46;
 Best Local Similarity 80.0%; Pred. No. 1.2e+04;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccgcgggtcacagtgtgcg 20
 ||||| ||||| ||||| |||||

Db 41 CCGCCGCTCAATGTGGTCG 22

RESULT 3

AU107932

LOCUS AU107932 50 bp mRNA EST 05-APR-2001
 DEFINITION ZRV62008, mRNA sequence.

ACCESSION AU107932

VERSION AU107932.1 GI:13557454

KEYWORDS EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

AUTHORS Suzuki.Y., Tsunoda.T., Taira.H., Mizushima-Sugano.J., Sese,J., Hata
 .H., Ota.T., Isoqai.T., Tanaka.T., Nakamura.Y., Morishita,S., Okubo
 .K., Suyama,A. and Sugano,S.

TITLE Fine Structural analysis of transcription start sites of human

JOURNAL

COMMENT

Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp
 Suzuki.Y., Yoshitomo-Nakagawa.K., Maruyama.K., Suyama,A. and Sugano
 .S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

Source

1..50

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="ZRV62008"

/clone_lib="Sugano Homo sapiens cDNA library"

6 a 19 c 15 g 9 t 1 others

BASE COUNT

ORIGIN

Query Match 66.0%; Score 13.2; DB 107; Length 50;
 Best Local Similarity 83.3%; Pred. No. 1.9e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ccgcgggtcacagtgtgc 19

||||| ||||| ||||| |||||

Db 17 CTCGGCGCCCAAGTGGTC 34

RESULT 4

AZ453043/c

LOCUS AZ453043 40 bp DNA GSS 04-OCT-2000

DEFINITION IM0254A13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0254A13 F, DNA sequence.

ACCESSION AZ453043

VERSION AZ453043.1 GI:10610442

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 40)

REFERENCE

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 .M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

JOURNAL

COMMENT

Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0254 row: A column: 13

Seq primer: CGTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 40.

FEATURES

Source

1..40

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0254A13"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 12 a 11 c 9 g 8 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 244; Length 40;
 Best Local Similarity 87.5%; Pred. No. 3e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ccgcgggtcacagtgg 17

||||| ||||| ||||| |||||

Db 26 CGGGGCGCTCACAGTGG 11

RESULT 5

AA822940/c

LOCUS AA822940 41 bp mRNA EST 17-FEB-1998

DEFINITION vp30g03.r1 Barstead mouse proximal colon MPLR86 Mus musculus cDNA
 clone IMAGE:1078228 5' similar to gb:U13705 Mus musculus domesticus

C57BL/6J plasma glutathione (MOUSE);, mRNA sequence.

ACCESSION AA822940

VERSION AA822940.1 GI:2892808

KEYWORDS	EST.	JOURNAL
SOURCE	house mouse,	COMMENT

EST.
house mouse.
Mus musculus
Eumariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 41)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston.R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:592524

```

Trace considered overall poor quality
Seq primer: -28mi3 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .41
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:1078228"
/cloned_lib="Barstead mouse proximal colon MPLRB6"
/dev_stage="7 day juvenile"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACGATCTGAGTGGGACGCCGCCCTTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[AAATCGCATCCTTG], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead. "
11 a 12 c 12 g 6 t

```

```

BASE COUNT      11 a 12 c 12 g 6 t
ORIGIN
Query Match      64.0%; Score 12.8; DB 12; Length 41;
Best Local Similarity 87.5%; Pred. No. 3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Query Match	64.0%;	Score 12.8;	DB 12;	Length 41;
Best Local Similarity	87.5%;	Pred. No. 3e+04;		
Matches 14;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	4	cagggtccacagtgtgc	19	
Db	27	CGCGCTCAGTGTGC	12	
RESULT	6			
LOCUS	AI762402	25 bp	mrna	EST
DEFINITION	wh65e10.x1 NCI_CGAP_kid11 Homo sapiens			cDNA clone IMAGE:2385642 3'
	similar to TR:P78421 P78421 CHROMOSOME 16P13			BAC CLONE
	CIT987SK-962B4 COMPLETE SEQUENCE; HTGS			PHASE 3 ; mRNA sequence.

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cdapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmer-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio-llnl.gov/hbrp/image/image.html

```
Trace considered overall poor quality
Insert Length: 436   Std error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1..225
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2385642"
/clone_lib="NCI_CGAP_Kid11"
/lab_host="DH10B"
/notice="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323376-1323911, 1456007-1456775, and
1500952-1502855). Subsequently by Bento Soares and M.
```

```

BASE COUNT      3 a          7 c          10 g          5 t
ORIGIN

Query Match      63.0%; Score 12.6; DB 24; Length 25;
Best Local Similarity 78.9%; Pred. No. 3.7e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

RESULT	7					
LOCUS	AA845818/c					
DEFINITION		AA845818	31 bp	mRNA	EST	04-MAR-1998
		ak85f01.s1 Barstead spleen HPLRB2 Homo sapiens cDNA clone				
		IMAGE:I414681 3' similar to SW:CIQC_HUMAN P02747 COMPLEMENT C1Q				
		SUBCOMPONENT, C CHAIN PRECURSOR. ; mRNA sequence.				

SOURCE	Humani.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 31) Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, N., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
TITLE	WasRU-NCI human EST Project
JOURNAL	Unpublished (1997)
COMMENT	Contact: Wilson RK

TITLE	WasRU-MCI human EST Project
JOURNAL	Unpublished (1997)
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LBNL : contact the


```

/db_xref="taxon:10090"
/clone="UUC2M0117121"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g114732114/gb1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      14 a      10 c      23 g      3 t
ORIGIN

Query Match      63.0%; Score 12.6; DB 250; Length 50;
Best Local Similarity 78.9%; Pred. No. 3.8e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ccgcgggctcacagtggctc 19
   | ||||| ||||| |||||
Db 30 CGCGGGCTCACTCGGAC 12

RESULT 10
TA233A12Q
LOCUS
DEFINITION
T. brucei sheared genomic DNA clone 233a12, reverse sequence, genomic survey sequence.
ACCESSION
AL481198
VERSION
AL481198.1 GI:11846892
KEYWORDS
GSS.
SOURCE
Trypanosoma brucei.
ORGANISM
Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE
1 (bases 1 to 44)
AUTHORS
Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.
Direct Submission
Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nhlesanger.ac.uk
COMMENT
Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nhlesanger@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES
Location/Qualifiers
source
1..44
/organism="Trypanosoma brucei"

```

```

/strain="TREU927"
/db_xref="taxon:5691"
/clone="233a12"
BASE COUNT      12 a      10 c      12 g      10 t
ORIGIN

Query Match      60.0%; Score 12; DB 258; Length 44;
Best Local Similarity 75.0%; Pred. No. 7.5e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ccgcgggctcacagtggctc 20
   | ||||| ||||| |||||
Db 14 CAGCGGGATCACAGGTCG 33

RESULT 11
AU106948/c
LOCUS
DEFINITION
AU106948 Sugano Homo sapiens cDNA library Homo sapiens CDNA clone
CAS03735, mRNA sequence.
ACCESSION
AU106948
VERSION
AU106948.1 GI:13556469
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 50)
AUTHORS
Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo, K., Suyama, A. and Sugano, S.
TITLE
Fine structural analysis of transcription start sites of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL
Unpublished (2001)
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
Location/Qualifiers
source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CAS03735"
BASE COUNT      6 a      16 c      24 g      4 t
ORIGIN

Query Match      60.0%; Score 12; DB 107; Length 50;
Best Local Similarity 75.0%; Pred. No. 7.5e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ccgcgggctcacagtggctc 20
   | ||||| ||||| |||||
Db 36 CCGCGGGCGCGCAGCAGCTCG 17

RESULT 12
TA140A04P/c
LOCUS
DEFINITION
T. brucei sheared genomic DNA clone 140a04, forward sequence, genomic survey sequence.
ACCESSION
AL466405
VERSION
AL466405.1 GI:11835760
KEYWORDS
GSS.
SOURCE
Trypanosoma brucei.
ORGANISM
Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

```


/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
/note="Organ: embryo; Vector: pSPORT; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally from mRNA prepared
from 800 blastocysts. Primer: SalI(df): cDNAs were
5'-CGGTGCGACGCGACCGTCTTTT-3'. cDNAs were
cloned into the NotI/SalI sites of a pSPORT vector (Life
Technologies). Two different size selections: B1 (larger
inserts) and B3."
BASE COUNT 10 a 9 c 12 g 6 t
ORIGIN

Query Match 58.0%; Score 11.6; DB 7; Length 37;
Best Local Similarity 77.8%; Pred. No. 1.2e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ccgcgggctcacagtgt 18
||| ||||| |||
Db 37 CCGTGGGCTCAGGT 20

RESULT 15
AA607492/c
LOCUS
DEFINITION
AA607492 46 bp mRNA EST 30-SEP-1997
vm61h02.r1 Stratagene mouse Tcell 937311 Mus musculus cDNA clone
IMAGE:1002771 5' similar to SW:RS9_RAT P29314 40S RIBOSOMAL PROTEIN
S9. mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AA607492.1 GI:2456385
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 46)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:566987

possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 1.
FEATURES
Location/Qualifiers
1..46
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1002771"
/clone_lib="Stratagene mouse Tcell 937311"
/tissue_type="Tcell"
/dev_stage="M30 CD4+ cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: blood; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. M30 CD4+ cells. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3'."

BASE COUNT 10 a 13 c 14 g
ORIGIN

Query Match 58.0%; Score 11.6; DB 9; Length 46;
Best Local Similarity 77.8%; Pred. No. 1.2e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ccgcgggctcacagtgt 18
||| ||||| |||
Db 46 CCGTGGGCTCAGGT 29

Search completed: October 2, 2001, 15:00:45
Job time: 10814 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:35 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-4
Perfect score: 20
Sequence: 1 ggcgcgtgaagcgaaggca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_om:*

8: gb_ov:*

9: gb_pat1:*

10: gb_pat2:*

11: gb_ph:*

12: gb_pli:*

13: gb_pl2:*

14: gb_pl3:*

15: gb_pl4:*

16: em_ba1:*

17: em_ba2:*

18: em_fun:*

19: em_htgo_hum:*

20: em_htgo_inv:*

21: em_htgo_rod:*

22: em_htg_hum1:*

23: em_htg_hum2:*

24: em_htg_hum3:*

25: em_htg_hum4:*

26: em_htg_hum5:*

27: em_htg_hum6:*

28: em_htg_hum7:*

29: em_htg_hum8:*

30: em_htg_inv1:*

31: em_htg_inv2:*

32: em_htg_other:*

33: em_htg_rod:*

34: em_hum1:*

35: em_hum2:*

36: em_hum3:*

37: em_hum4:*

38: em_hum5:*

39: em_hum6:*

40: em_hum7:*

41: em_in:*

42: em_om:*

43: em_or:*

44: em_ov:*

45: em_pat:*

46: em_ph:*

47: em_pl:*

48: em_ro:*

49: em_sts:*

50: em_sy:*

51: em_un:*

52: em_vi:*

53: gb_sts1:*

54: gb_sts2:*

55: gb_sts3:*

56: gb_sy:*

57: gb_un:*

58: gb_vil:*

59: gb_vil2:*

60: gb_htg1:*

61: gb_htg2:*

62: gb_htg3:*

63: gb_htg4:*

64: gb_htg5:*

65: gb_htg6:*

66: gb_htg7:*

67: gb_htg8:*

68: gb_htg9:*

69: gb_htg10:*

70: gb_htg11:*

71: gb_htg12:*

72: gb_htg13:*

73: gb_htg14:*

74: gb_htg15:*

75: gb_htg16:*

76: gb_htg17:*

77: gb_htg18:*

78: gb_htg19:*

79: gb_htg20:*

80: gb_htg21:*

81: gb_htg22:*

82: gb_htg23:*

83: gb_htg24:*

84: gb_htg25:*

85: gb_pr1:*

86: gb_pr2:*

87: gb_pr3:*

88: gb_pr4:*

89: gb_pr5:*

90: gb_pr6:*

91: gb_pr7:*

92: gb_pr8:*

93: gb_pr9:*

94: gb_rol:*

95: gb_rol2:*

96: gb_in4:*

97: gb_pr10:*

98: em_ba3:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	13.2	66.0	21	10	I27712
2	13.2	66.0	21	10	I83056
3	12.4	62.0	21	10	AX094895
4	12.2	61.0	22	9	A09940
5	12.2	61.0	38	9	A47740
6	12.2	61.0	38	9	AR019411
7	12.2	61.0	47	97	HUMTCVD1DK
8	12.2	61.0	50	9	AR092516
					I27712 Sequence 19
					I83056 Sequence 19
					AX094895 Sequence
					A09940 oligonucleo
					A47740 Sequence 4
					AR019411 Sequence
					L32460 Human (clon
					AR092516 Sequence

c 9 12 60.0 20 10 121584 I21584 Sequence 4
 10 12 60.0 48 9 A60848 A60848 Sequence 15
 11 11.8 59.0 20 9 AX020035 AX020035 Sequence
 12 11.6 58.0 25 10 AX076946 AX076946 Sequence
 13 11.6 58.0 27 9 AR044492 AR044492 Sequence
 14 11.6 58.0 36 9 A93959 A93959 Sequence 13
 15 11.6 58.0 43 9 AR012355 AR012355 Sequence
 c 16 11.4 57.0 23 10 E30097 E30097 Endo-beta-N
 17 11.4 57.0 24 3 S58583 S58583 trNA(GGCAla
 18 11.4 57.0 29 10 E12798 E12798 Primer for
 19 11.2 56.0 28 10 I71874 I71874 Sequence 72
 20 11.2 56.0 30 9 A78515 A78515 Sequence 6
 21 11.2 56.0 30 10 I08716 I08716 Sequence 4
 22 11.2 56.0 30 10 I08721 I08721 Sequence 9
 23 11.2 56.0 30 10 I55881 I55881 Sequence 11
 24 11.2 56.0 33 9 AR004362 AR004362 Sequence
 25 11.2 56.0 33 9 AR097153 AR097153 Sequence
 26 11.2 56.0 33 10 I82837 I82837 Sequence 16
 27 11.2 56.0 35 10 E30450 E30450 Impartation
 28 11.2 56.0 37 9 AR078298 AR078298 Sequence
 29 11.2 56.0 38 9 AR050714 AR050714 Sequence
 30 11.2 56.0 38 9 AR093351 AR093351 Sequence
 31 11.2 56.0 41 9 AR029572 AR029572 Sequence
 32 11.2 56.0 41 9 AR098525 AR098525 Sequence 87
 33 11.2 56.0 41 10 I41485 I41485 Sequence
 34 11.2 56.0 50 94 MMU41904 U41904 Mus musculus
 35 11.2 56.0 50 94 MMU41940 U41940 Mus musculus
 36 11.2 56.0 50 94 MMU41956 U41956 Mus musculus
 37 11.2 56.0 50 94 MMU41971 U41971 Mus musculus
 38 11 55.0 21 45 E10439 E10439 Primer, 9/2
 39 11 55.0 22 9 AR024068 AR024068 Sequence
 40 11 55.0 22 94 MMTCL1902A MMTCL1902A
 41 11 55.0 24 9 AR089906 AR089906 Sequence
 42 11 55.0 25 10 I92340 I92340 Sequence 1
 43 11 55.0 25 10 I92341 I92341 Sequence 2
 44 11 55.0 26 10 I56539 I56539 Sequence 12
 45 11 55.0 27 97 S41969 S41969 p53-p53 onc

ALIGNMENTS

RESULT 1
 I27712
 LOCUS I27712 21 bp DNA PAT 06-FEB-1997
 DEFINITION Sequence 19 from patent US 5567583.
 ACCESSION I27712
 VERSION I27712.1 GI:1818488
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 21)
 AUTHORS Wang,C.J., and Wu,K.
 TITLE Methods for reducing non-specific priming in DNA detection
 JOURNAL Patent: US 5567583-A 19 22-OCT-1996;
 FEATURES Location/Qualifiers
 source 1..21
 BASE COUNT 2 a 6 c 11 g 2 t
 ORIGIN
 Query Match 66.0%; Score 13.2; DB 10; Length 21;
 Best Local Similarity 83.3%; Pred. No. 8.8e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 gcgcctgaagcgaagc 19
 ||| ||||| |||
 Db 1 GCGGCGTGAAGCGCGC 18
 ||| ||||| |||
 RESULT 2
 I27712
 LOCUS I27712 21 bp DNA PAT 10-JUN-1998
 DEFINITION Sequence 19 from patent US 5712386.
 ACCESSION I27712
 VERSION I27712.1 GI:3211353
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 21)
 AUTHORS Wang,C.J., and Wu,K.
 TITLE Kits for detecting a target nucleic acid with blocking
 oligonucleotides
 JOURNAL Patent: US 5712386-A 19 27-JAN-1998;
 FEATURES Location/Qualifiers
 source 1..21
 BASE COUNT 2 a 6 c 11 g 2 t
 ORIGIN
 Query Match 66.0%; Score 13.2; DB 10; Length 21;
 Best Local Similarity 83.3%; Pred. No. 8.8e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 gcgcctgaagcgaagc 19
 ||| ||||| |||
 Db 1 GCGGCGTGAAGCGCGC 18
 ||| ||||| |||
 RESULT 4
 A09940/c
 LOCUS A09940 22 bp DNA PAT 25-JAN-1994
 DEFINITION oligonucleotide.
 ACCESSION A09940
 VERSION A09940.1 GI:492306
 KEYWORDS
 SOURCE unidentified.

```
ORGANISM unidentified
REFERENCE 1 (bases 1 to 22)
AUTHORS Moureau,P., Derclaye,I., Delor,I. and Cornelis,G.
TITLE Nucleic acid probes useful for detecting specifically different
        bacterial species of the genus Campylobacter
JOURNAL Patent: EP 0350392-A 2 10-JAN-1990;
FEATURES   IRE-MEDGENIX S.A
SOURCE     Location/Qualifiers
           1..22
           /organism="unidentified"
           /db_xref="taxon:32644"
BASE COUNT 2 a 7 c 6 g 7 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 22;
Best Local Similarity 82.4%; Pred. No. 2.7e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gccgcgtgaagcgaagg 18
   ||||| ||||| |||||
Db 17 GCTCCGTGAACCAAGG 1

RESULT 5
LOCUS A47740 38 bp DNA PAT 07-MAR-1997
DEFINITION Sequence 4 from Patent W09533833.
ACCESSION A47740
VERSION A47740.1 GI:2301650
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 38)
AUTHORS Wood,P.C. and Quirk,A.V.
TITLE YEAST STRAINS
JOURNAL Patent: WO 9533833-A 4 14-DEC-1995;
COMMENT DELTA BIOTECHNOLOGY LTD (GB)
FEATURES Other publication AU 2626295 960104.
SOURCE Location/Qualifiers
           1..38
           /organism="unidentified"
           /db_xref="taxon:32644"
BASE COUNT 7 a 9 c 13 g 9 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 38;
Best Local Similarity 82.4%; Pred. No. 2.4e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gccgtgaagcgaagcca 20
   ||||| ||||| |||||
Db 1 GCCGTGTAGCGAGGAA 17

RESULT 6
LOCUS AR019411 38 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 4 from patent US 5783423.
ACCESSION AR019411
VERSION AR019411.1 GI:3974525
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 38)
AUTHORS Wood,P.Carolyn and Quirk,A.Victor.
TITLE Yeast strains
JOURNAL Patent: US 5783423-A 4 21-JUL-1998;
FEATURES Location/Qualifiers

ORGANISM unidentified
SOURCE /organism="unknown"
BASE COUNT 10 a 18 c 18 g 4 t
ORIGIN

ORGANISM unidentified
SOURCE /organism="unknown"
BASE COUNT 7 a 9 c 13 g 9 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 38;
Best Local Similarity 82.4%; Pred. No. 2.4e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gccgtgaagcgaagcca 20
   ||||| ||||| |||||
Db 1 GCCGTGTAGCGAGGAA 17

RESULT 7
LOCUS HUMTCVD1DK/c 47 bp mRNA PRI 10-FEB-1995
DEFINITION Human (clone: 3cpj20) T-cell receptor delta-chain (V-delta-1) mRNA.
ACCESSION L32460
VERSION L32460.1 GI:497530
KEYWORDS T-cell receptor; delta chain.
SOURCE Homo sapiens intestine cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 47)
AUTHORS Chowers,Y., Holtmeier,W., Harwood,J., Morzycka-Wroblewska,E. and
        Kagnoff,M.F.
TITLE The V delta 1 T cell receptor repertoire in human small intestine
        and colon
JOURNAL J. Exp. Med. 180 (1), 183-190 (1994)
MEDLINE 94275371
FEATURES Location/Qualifiers
SOURCE 1..47
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="intestine"
BASE COUNT 9 a 15 c 9 g 14 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 97; Length 47;
Best Local Similarity 82.4%; Pred. No. 2.3e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gccgtgaagcgaagcca 20
   ||||| ||||| |||||
Db 28 GCCATGTAGCGTAGGCA 12

RESULT 8
LOCUS AR092516 50 bp DNA PAT 08-SEP-2000
DEFINITION Sequence 46 from patent US 5998168.
ACCESSION AR092516
VERSION AR092516.1 GI:10019270
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 50)
AUTHORS Loomore,S.M., Yacoub,R.Khayyam, Zealey,G.Ross and Klein,M.Henri.
TITLE Expression of gene products from genetically manipulated strains of
        bordetella
JOURNAL Patent: US 5998168-A 46 07-DEC-1999;
FEATURES Location/Qualifiers
SOURCE 1..50
/organism="unknown"
BASE COUNT 10 a 18 c 18 g 4 t
ORIGIN
```

Query Match 61.0%; Score 12.2; DB 9; Length 50;
 Best Local Similarity 82.4%; Pred. No. 2.3e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcgaag 17
 Db 23 GGCgcgtgacgcgaag 39

RESULT 9
 I21584/c
 LOCUS I21584 20 bp DNA PAT 07-OCT-1996
 DEFINITION Sequence 4 from patent US 5521301.
 ACCESSION I21584
 VERSION I21584.1 GI:1601938
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 20)
 AUTHORS Wallace,R.Bruce and Ugozzoli,L.
 TITLE Genotyping of multiple allele systems
 JOURNAL Patent: US 5521301-A 4 28-MAY-1996;
 FEATURES Location/Qualifiers
 source 1..20
 /organism="unknown"
 BASE COUNT 2 a 7 c 7 g 4 t
 ORIGIN

Query Match 60.0%; Score 12; DB 10; Length 20;
 Best Local Similarity 75.0%; Pred. No. 3.4e+05;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcgaagca 20
 Db 20 GGCgcctacagcgctgcca 1

RESULT 10
 A60848
 LOCUS A60848 48 bp DNA PAT 06-MAR-1998
 DEFINITION Sequence 157 from Patent WO9708320.
 ACCESSION A60848
 VERSION A60848.1 GI:3715469
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 48)
 AUTHORS Knappik,A., Pack,P., Ilag,V., Ge,L., Moroney,S. and Plueckthun,A.
 TITLE PROTEIN/(POLY)PEPTIDE LIBRARIES
 JOURNAL Patent: WO 9708320-A 157 06-MAR-1997;
 MORPHOSYS PROTEINOPTIMIERUNG (DE)
 FEATURES Location/Qualifiers
 source 1..48
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 9 a 14 c 17 g 8 t
 ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 48;
 Best Local Similarity 75.0%; Pred. No. 2.9e+05;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcgaagca 20
 Db 1 GCGcgttagagcccaagcca 20

RESULT 11
 AX020035

LOCUS AX020035 20 bp DNA PAT 07-SEP-2000
 DEFINITION Sequence 49 from Patent WO9937764.
 ACCESSION AX020035
 VERSION AX020035.1 GI:10043864
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 20)
 AUTHORS Veugelers,M.P. and David,G.J.
 TITLE New members of the glypican gene family
 JOURNAL Patent: WO 9937764-A 49 29-JUL-1999;
 VEUGELERS MARK PAUL DITMAR (BE); VLAAMS INTERUNIV INST BIOTECH
 (BE); DAVID GUIDO JOSEPH FRANS (BE)
 FEATURES Location/Qualifiers
 source 1..20
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 9 a 3 c 6 g 2 t
 ORIGIN

Query Match 59.0%; Score 11.8; DB 9; Length 20;
 Best Local Similarity 86.7%; Pred. No. 4.2e+05;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 cgtgaagcgaagcca 20
 Db 5 CGTGAAGCAAGACA 19

RESULT 12
 AX076946
 LOCUS AX076946 25 bp DNA PAT 22-FEB-2001
 DEFINITION Sequence 58 from Patent WO0105836.
 ACCESSION AX076946
 VERSION AX076946.1 GI:13121599
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM artificial sequence.
 REFERENCE 1 (bases 1 to 25)
 AUTHORS Botstein,D., Goddard,A., Gurney,A.L., Hillan,K.J., Roy,M.A. and Wood,W.I.
 TITLE Polypeptidic compositions and methods for the treatment of tumors
 JOURNAL Patent: WO 0105836-A 58 25-JAN-2001;
 Genentech, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..25
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Synthetic Oligonucleotide Probe"
 BASE COUNT 9 a 8 c 5 g 3 t
 ORIGIN

Query Match 58.0%; Score 11.6; DB 10; Length 25;
 Best Local Similarity 77.8%; Pred. No. 5.1e+05;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ggcgcgtgaagcgaagc 19
 Db 3 GCGCGCAAGCAAGAC 20

RESULT 13
 AR044492
 LOCUS AR044492 27 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 32 from patent US 5817495.
 ACCESSION AR044492
 VERSION AR044492.1 GI:5965957
 KEYWORDS

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 27)
AUTHORS Pedersen,A.Hjelholt, Vind,J., Svendsen,A., Cherry,J.R., Lamsa,M.,
Schneider,P. and Jensen,B.Rostgaard.
TITLE H.sub.2 O.sub.2 -stable peroxidase variants
JOURNAL Patent: US 5817495-A 32 06-OCT-1998;
FEATURES Location/Qualifiers
source 1..27
BASE COUNT 5 a 9 c 11 g 2 t
ORIGIN /organism="unknown"

Query Match 58.0%; Score 11.6; DB 9; Length 27;
Best Local Similarity 77.8%; Pred. No. 5e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggcgcctgaagcgaagg 18
||||| ||| |||||
Db 10 GGC CGCGCGAGAGG 27

RESULT 14
A93959/c
LOCUS A93959 36 bp DNA PAT 22-JAN-2000
DEFINITION Sequence 13 from Patent WO9718322.
ACCESSION A93959
VERSION A93959.1 GI:6742060
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 36)
AUTHORS Dahm,M.W.
TITLE METHOD OF QUANTIFYING TUMOUR CELLS IN A BODY FLUID AND A SUITABLE
TEST KIT
JOURNAL Patent: WO 9718322-A 13 22-MAY-1997;
DAHM MICHAEL W (DE)
FEATURES Location/Qualifiers
source 1..36
BASE COUNT 7 a 15 c 8 g 6 t
ORIGIN /organism="unidentified"
/db_xref="taxon:32644"

Query Match 58.0%; Score 11.6; DB 9; Length 36;
Best Local Similarity 77.8%; Pred. No. 4.7e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggcgcctgaagcgaagg 18
||||| ||| |||||
Db 18 GGCACGCGAATCGATCG 1

RESULT 15
AR012355
LOCUS AR012355 43 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 46 from patent US 5763284.
ACCESSION AR012355
VERSION AR012355.1 GI:3970345
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 43)
AUTHORS Tal,R., Wong,H.C., Casipit,C., Chavallaz,P. and Wittman,V.
TITLE Methods for peptide synthesis and purification
JOURNAL Patent: US 5763284-A 46 09-JUN-1998;
FEATURES Location/Qualifiers
source 1..43

BASE COUNT 8 a 11 c 15 g 9 t
ORIGIN /organism="unknown"

Query Match 58.0%; Score 11.6; DB 9; Length 43;
Best Local Similarity 77.8%; Pred. No. 4.6e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggcgcctgaagcgaagg 18
||||| ||| |||||
Db 1 GGGCCCATGGATCGAGG 18

Search completed: October 2, 2001, 15:56:37
Job time: 14161 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:00:45 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-4

Perfect score: 20

Sequence: 1 ggcgcgtgaagcgaagca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: gb_estl1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est25.*
26: gb_est26.*
27: gb_est27.*
28: gb_est28.*
29: gb_est29.*
30: gb_est30.*
31: gb_est31.*
32: gb_est32.*
33: gb_est33.*
34: gb_est34.*
35: gb_est35.*
36: gb_est36.*
37: gb_est37.*
38: gb_est38.*
39: gb_est39.*
40: gb_est40.*
41: gb_est41.*
42: gb_est42.*
43: gb_est43.*
44: gb_est44.*
45: gb_est45.*
46: gb_est46.*
47: gb_est47.*
48: gb_est48.*
49: gb_est49.*
50: gb_est50.*
51: gb_est51.*
52: gb_est52.*
53: gb_est53.*
54: gb_est54.*
55: gb_est55.*
56: gb_est56.*
57: gb_est57.*
58: gb_est58.*
59: gb_est59.*
60: gb_est60.*
61: gb_est61.*
62: gb_est62.*
63: gb_est63.*
64: gb_est64.*
65: gb_est65.*
66: gb_est66.*
67: gb_est67.*
68: gb_est68.*
69: gb_est69.*
70: gb_est70.*
71: gb_est71.*
72: gb_est72.*
73: gb_est73.*
74: gb_est74.*
75: gb_est75.*
76: gb_est76.*
77: gb_est77.*
78: gb_est78.*
79: gb_est79.*
80: gb_est80.*
81: gb_est81.*
82: gb_est82.*
83: gb_est83.*
84: gb_est84.*
85: gb_est85.*
86: gb_est86.*
87: gb_est87.*
88: gb_est88.*
89: gb_est89.*
90: gb_est90.*
91: gb_est91.*
92: gb_est92.*
93: gb_est93.*
94: gb_est94.*
95: gb_est95.*
96: gb_est96.*
97: gb_est97.*
98: gb_est98.*
99: gb_est99.*
100: gb_est100.*
101: gb_est101.*
102: gb_est102.*
103: gb_est103.*
104: gb_est104.*
105: gb_est105.*
106: gb_est106.*
107: gb_est107.*
108: gb_est108.*
109: gb_est109.*
110: gb_est110.*
111: gb_est111.*
112: gb_est112.*
113: gb_est113.*
114: gb_est114.*
115: gb_est115.*
116: gb_est116.*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	14	70.0	50	107	AU103598
C 2	13.6	68.0	50	107	AU105195
C 3	13.6	68.0	50	107	AU105195
C 4	13.6	68.0	50	107	AU105197
C 5	13.2	66.0	47	250	AU105198
C 6	13.2	66.0	50	107	AU105198
C 7	12.8	64.0	50	107	AZ816650
C 8	12.8	64.0	50	107	2M085003
C 9	12.8	64.0	50	107	AU105437
C 10	12.8	64.0	50	107	AU105437
C 11	12.8	64.0	50	107	AU103765
C 12	12.8	64.0	50	107	AU103765
C 13	12.8	64.0	50	107	AU103768
C 14	12.8	64.0	50	107	AU103771
C 15	12.8	64.0	50	107	AU103771
C 16	12.8	64.0	50	107	AU103787
C 17	12.8	64.0	50	107	AU103787
C 18	12.6	63.0	50	107	AU103791
C 19	12.4	62.0	50	107	AU103791
C 20	12.2	61.0	50	107	AU103974
C 21	12.2	61.0	50	107	AU103975
C 22	12.2	61.0	50	107	AU103975
C 23	12.2	61.0	50	107	AU103977
C 24	12.2	61.0	50	107	AU103977
C 25	12.2	61.0	50	107	AU103977
C 26	12.2	61.0	50	107	AU103977
C 27	12.2	61.0	50	107	AU103977
C 28	12.2	61.0	50	107	AU103977
C 29	12.2	61.0	50	107	AU103977
C 30	12.2	61.0	50	107	AU103977
C 31	11.8	59.0	46	14	AU103980
C 32	11.8	59.0	50	107	AU103980
C 33	11.8	59.0	50	107	AU103980
C 34	11.8	59.0	50	107	AU103980
C 35	11.8	59.0	50	107	AU103980
C 36	11.8	59.0	50	107	AU103980
C 37	11.8	59.0	50	107	AU103980
C 38	11.8	59.0	50	107	AU103980
C 39	11.8	59.0	50	107	AU103980
C 40	11.8	59.0	50	107	AU103980
C 41	11.8	59.0	50	107	AU103980
C 42	11.8	59.0	50	107	AU103980
C 43	11.8	59.0	50	107	AU103980
C 44	11.8	59.0	50	107	AU103980
C 45	11.8	59.0	50	107	AU103980

ALIGNMENTS

RESULT 1
 AU103598/c 50 bp mRNA EST 05-APR-2001
 LOCUS AU103598 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 DEFINITION COLF6124, mRNA sequence.
 ACCESSION AU103598
 VERSION AU103598.1 GI:13553119
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
 ,K., Suyama,A. and Sugano,S.
 TITLE Fine Structural analysis of transcription start sites of human

JOURNAL Unpublished (2001)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 ,S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
 source 1..50
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COLF6124"
 /clone_lib="Sugano Homo sapiens cDNA library"
 BASE COUNT 8 a 16 c 19 g 7 t
 ORIGIN

Query Match 70.0% Score 14; DB 107; Length 50;
 Best Local Similarity 100.0%; Pred. No. 6e+03;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggcgcgtgaagcga 15
 |||||
 Db 18 GCGCGTGAAGCGA 5

RESULT 2
 AU105195 50 bp mRNA EST 05-APR-2001
 LOCUS AU105195 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 DEFINITION HRC06645, mRNA sequence.
 ACCESSION AU105195
 VERSION AU105195.1 GI:13554716
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
 ,K., Suyama,A. and Sugano,S.
 TITLE Fine Structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries
 JOURNAL Unpublished (2001)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 ,S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
 source 1..50
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HRC06645"
 /clone_lib="Sugano Homo sapiens cDNA library"
 BASE COUNT 8 a 13 c 24 g 5 t
 ORIGIN

Query Match 68.0% Score 13.6; DB 107; Length 50;
 Best Local Similarity 80.0%; Pred. No. 9.5e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggcgcgtgaagcgaagcga 20
 |||||
 Db 1 GGAGCTGTGAGGCGCAGGCA 20

10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114.gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptorized mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 8 a 7 c 22 g 10 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 250; Length 47;
Best Local Similarity 83.3%; Pred. No. 1.5e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 cgcgcgtgaagcgaagcga 20
| | | | | | | | | | | | | | | | | | | | | |
Db 38 CCCCGTGAAGAGAAAGCA 21

RESULT 6
AUI05437/c

LOCUS AUI05437 50 bp mRNA EST 05-APR-2001
DEFINITION AUI05437 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP15178, mRNA sequence.

ACCESSION AUI05437
VERSION AUI05437
KEYWORDS EST.
SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Tanaka,H., Mizushima-Sugano,J., Sese,J., Hata
K., Suyama,A. and Sugano,S.

TITLE Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL Unpublished (2001)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yszukie@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES Location/Qualifiers

source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP15178"

BASE COUNT 2 a 21 c 15 g 12 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 107; Length 50;
Best Local Similarity 83.3%; Pred. No. 1.5e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggcgcctgaagcgaagcga 18
| | | | | | | | | | | | | | | | | | | | | |
Db 35 GGCCCCCGAAGCGAAG 18

RESULT 7
AUI03765/c

LOCUS AUI03765 50 bp mRNA EST 05-APR-2001
DEFINITION AUI03765 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP11908, mRNA sequence.

ACCESSION AUI03765
VERSION AUI03765.1 GI:13553286
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Tanaka,H., Mizushima-Sugano,J., Sese,J., Hata
K., Suyama,A. and Sugano,S.

TITLE Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL Unpublished (2001)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yszukie@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES Location/Qualifiers

source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP11908"

BASE COUNT 11 a 14 c 12 g 13 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 107; Length 50;
Best Local Similarity 87.5%; Pred. No. 2.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggcgcctgaagcgaagcga 16
| | | | | | | | | | | | | | | | | | | | | |
Db 27 GGCGAGGGAAGCGAA 12

RESULT 8
AUI03768/c

LOCUS AUI03768 50 bp mRNA EST 05-APR-2001
DEFINITION AUI03768 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP12651, mRNA sequence.

ACCESSION AUI03768
VERSION AUI03768.1 GI:13553289
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Tanaka,H., Mizushima-Sugano,J., Sese,J., Hata
K., Suyama,A. and Sugano,S.

TITLE Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL Unpublished (2001)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yszukie@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES Location/Qualifiers

source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP12651"

```

/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      8 a 14 c 11 g 17 t
ORIGIN

Query Match      64.0%; Score 12.8; DB 107; Length 50;
Best Local Similarity 87.5%; Pred. No. 2.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggccgcgtgaagcgaa 16
    ||||| | |||||
DB 39 GCGCGAGGGAAGCGAA 24

RESULT 9
AUI03771/c
LOCUS      AUI03771 50 bp mRNA EST 05-APR-2001
DEFINITION AUI03771 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP13114, mRNA sequence.
ACCESSION  AUI03771
VERSION    AUI03771.1 GI:13553292
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 50)
AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,K., Suyama,A. and Sugano,S.
            H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,S.
TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL   Unpublished (2001)
COMMENT   Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: yszuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             source
            1..50
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="HEP13114"
            /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      6 a 14 c 12 g 18 t
ORIGIN

Query Match      64.0%; Score 12.8; DB 107; Length 50;
Best Local Similarity 87.5%; Pred. No. 2.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggccgcgtgaagcgaa 16
    ||||| | |||||
DB 47 GCGCGAGGGAAGCGAA 32

RESULT 10
AUI03787/c
LOCUS      AUI03787 50 bp mRNA EST 05-APR-2001
DEFINITION AUI03787 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP15173, mRNA sequence.
ACCESSION  AUI03787
VERSION    AUI03787.1 GI:13553308
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 50)

```

```

AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL   Unpublished (2001)
COMMENT   Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: yszuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             source
            1..50
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="HEP15173"
            /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      6 a 13 c 11 g 20 t
ORIGIN

Query Match      64.0%; Score 12.8; DB 107; Length 50;
Best Local Similarity 87.5%; Pred. No. 2.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggccgcgtgaagcgaa 16
    ||||| | |||||
DB 47 GCGCGAGGGAAGCGAA 32

RESULT 11
AUI03791/c
LOCUS      AUI03791 50 bp mRNA EST 05-APR-2001
DEFINITION AUI03791 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP16708, mRNA sequence.
ACCESSION  AUI03791
VERSION    AUI03791.1 GI:13553312
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 50)
AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
TITLE     Fine Structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL   Unpublished (2001)
COMMENT   Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: yszuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             source
            1..50
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="HEP16708"
            /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      6 a 13 c 13 g 18 t
ORIGIN

Query Match      64.0%; Score 12.8; DB 107; Length 50;
Best Local Similarity 87.5%; Pred. No. 2.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ggcgcgtgaagcgaa 16
||||| |||||
Db 43 GCGCCAGGAAGCGAA 28

RESULT 12
AU103974/c

LOCUS AU103974 50 bp mRNA EST 05-APR-2001
DEFINITION AU103974 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP00461, mRNA sequence.

ACCESSION AU103974
VERSION AU103974.1 GI:13553495
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.

TITLE

Fine Structural analysis of transcription start sites of human

JOURNAL

Unpublished (2001)

COMMENT

Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP00461"
/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 6 a 20 c 14 g 10 t

Query Match

Best Local Similarity 64.0%; Score 12.8; DB 107; Length 50;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ggcgcgtgaagcgaa 17
||||| |||||
Db 25 GCGCCGGAAGCGAGG 10

RESULT 13
AU103975/c

LOCUS AU103975 50 bp mRNA EST 05-APR-2001
DEFINITION AU103975 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP10657, mRNA sequence.

ACCESSION AU103975
VERSION AU103975.1 GI:13553496
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.

TITLE

Fine Structural analysis of transcription start sites of human

JOURNAL

Unpublished (2001)

COMMENT

Contact: Yutaka Suzuki
Department of Virology

Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP10657"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 4 a 21 c 9 g 16 t
ORIGIN

Query Match

Best Local Similarity 64.0%; Score 12.8; DB 107; Length 50;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ggcgcgtgaagcgaa 17
||||| |||||
Db 46 GCGCCGGAAGCGAGG 31

RESULT 14
AU103977/c

LOCUS AU103977 50 bp mRNA EST 05-APR-2001
DEFINITION AU103977 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP17580, mRNA sequence.

ACCESSION AU103977
VERSION AU103977.1 GI:13553498
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.

TITLE

Fine Structural analysis of transcription start sites of human

JOURNAL

Unpublished (2001)

COMMENT

Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source
1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP17580"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 5 a 21 c 12 g 12 t
ORIGIN

Query Match

Best Local Similarity 64.0%; Score 12.8; DB 107; Length 50;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ggcgcgtgaagcgaa 17
||||| |||||
Db 28 GCGCCGGAAGCGAGG 13

RESULT 15
AU103980/c

LOCUS AU103980 50 bp mRNA EST 05-APR-2001
DEFINITION AU103980 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
KAT01573, mRNA sequence.
ACCESSION AU103980
VERSION AU103980.1 GI:13553501
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
K., Suyama,A. and Sugano,S.
TITLE Fine structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL Unpublished (2001)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
source
1..50
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="KAT01573"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 6 a 20 c 14 g 10 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 107; Length 50;
Best Local Similarity 87.5%; Pred. No. 2.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 gcgccgtgaagcgaag 17
||||| ||||| |
Db 25 GCGCCGGGAGCGGAGG 10

Search completed: October 2, 2001, 15:00:46
Job time: 10815 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:37 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-6
Perfect score: 20
Sequencing: 1 gaaactgcagaagcactga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_bal: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pil: *
13: gb_p12: *
14: gb_p13: *
15: gb_p14: *
16: em_bal: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_vi: *
53: gb_stsl: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_vil: *
59: gb_vil2: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_rol: *
95: gb_ro2: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	13.8	69.0	25	10	I26640
c 2	13.8	69.0	42	95	S77087
c 3	13.6	68.0	38	91	EPPROTII2
4	13.6	68.0	38	97	SEPROTII2
5	13.4	67.0	36	9	A69037
c 6	13.4	67.0	43	9	AR018055
c 7	13.4	67.0	43	9	AR071945
8	13.4	67.0	44	10	E05905
					I26640 Sequence 3
					S77087 T-cell rece
					AF195645 Erythroce
					AF195643 Semnopath
					A69037 Sequence 25
					AR018055 Sequence
					AR071945 Sequence
					E05905 Primer. 9/1

```
9 13.2 66.0 27 10 I27338 Sequence 13
10 13 E04313 PCR primer
11 12.8 64.0 22 9 A45762 Sequence 19
12 12.8 64.0 22 9 AR069418 Sequence 1
13 12.8 64.0 22 10 I43148
14 12.8 64.0 28 9 AR08097
15 12.8 64.0 28 9 AR08101
16 12.8 64.0 33 9 AR08100
17 12.8 64.0 34 9 A62490 Sequence 5
18 12.8 64.0 35 9 A69043 Sequence 1
19 12.8 64.0 47 9 HUMTCRVJ55
20 12.6 63.0 21 9 AR011280
21 12.6 63.0 21 10 I17918
22 12.6 63.0 30 9 AR066368
23 12.6 63.0 34 97 S80826
24 12.4 62.0 25 9 AR003624
25 12.4 62.0 33 9 AR08096
26 12.4 62.0 39 10 I79231
27 12.2 61.0 18 9 AR101074
28 12.2 61.0 18 10 AX068306
29 12.2 61.0 27 10 I33779
30 12.2 61.0 30 10 AX080135
31 12.2 61.0 31 9 AR078949
32 12.2 61.0 42 9 A36503
33 12.2 61.0 42 9 AR080136
34 12.2 61.0 42 10 AX097533
35 12.2 61.0 50 9 AR032784
36 12.2 61.0 50 10 I29524
37 12.2 61.0 50 10 I91198
38 12 60.0 23 9 A84872
39 12 60.0 23 9 AR089998
40 12 60.0 29 9 A03947
41 12 60.0 29 9 A15005
42 12 60.0 30 9 A46142
43 12 60.0 30 10 I79830
44 12 60.0 32 9 AR052426
45 12 60.0 32 9 AR082414 Sequence
```

ALIGNMENTS

```
RESULT 1
LOCUS I26640 25 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 3 from patent US 5559009.
ACCESSION I26640
VERSION I26640.1 GI:1606510
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 25)
AUTHORS Chandy,K.G., Kalman,K., Chandy,G. and Gutman,G.A.
TITLE Voltage-gated potassium channel gene, KV1.7, vectors and host cells comprising the same, and recombinant methods of making potassium channel proteins
JOURNAL Patent: US 5559009-A 3 24-SEP-1996;
FEATURES Location/Qualifiers
source 1..25
2 a 11 c 5 g 7 t
BASE COUNT 2 a 11 c 5 g 7 t
ORIGIN
```

Query Match 69.0%; Score 13.8; DB 10; Length 25;
Best Local Similarity 88.2%; Pred. No. 2.le+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 gaactgcagaaggc 17
Db 17 GGAAGTCAGAGGAC 1

```
RESULT 2
LOCUS S77087/c 42 bp mRNA ROD 14-AUG-2000
DEFINITION T-cell receptor beta chain V-J region [CDR3 region] [mice, transgenic, mRNA Partial, 42 nt].
ACCESSION S77087
VERSION S77087.1 GI:242717
KEYWORDS
SOURCE Mus sp. transgenic.
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 42)
AUTHORS Jorgensen,J.L., Esser,U., de St Groth,B.F., Reay,P.A. and Davis,M.M.
TITLE Mapping T-cell receptor-peptide contacts by variant peptide immunization of single-chain transgenics
JOURNAL Nature 355 (6357), 224-230 (1992)
MEDLINE 92114961
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 77087] from the original journal article. This sequence comes from FIG. 4 c.
FEATURES Location/Qualifiers
source 1..42
/organism="Mus sp."
/db_xref="taxon:10095"
<1..>42
/note="This sequence comes from FIG. 4 c; conceptual translation presented here differs from translation in publication"
/codon_start=1
/product="T-cell receptor beta chain V-J region"
/protein_id="AAB20964.2"
/db_xref="GI:9800659"
/translation="CASSAFWGVAEQFF"
BASE COUNT 5 a 10 c 13 g 14 t
ORIGIN
```

Query Match 69.0%; Score 13.8; DB 95; Length 42;
Best Local Similarity 88.2%; Pred. No. 2.le+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 aactgcagaaggcactg 19
Db 25 AACCCAGAGGCACTG 9

```
RESULT 3
LOCUS EPPROTI12 38 bp DNA PRI 23-DEC-2000
DEFINITION Erythrocebus patas protamine P2 gene, exon 2 and complete cds.
ACCESSION AF195645
VERSION AF195645.1 GI:11990583
KEYWORDS
SEGMENT 2 of 2
SOURCE red quonon.
ORGANISM Erythrocebus patas
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopithecinae; Erythrocebus. 1 (bases 1 to 38)
AUTHORS Karanth,P., Stewart,C.B., Holt,R.A., deKoning,J. and Messier,W.
TITLE Positive Darwinian selection on the lineage leading to humans
JOURNAL Unpublished
AUTHORS Karanth,P. and Stewart,C.B.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1999) Biological Sciences, State University of New York, 1400 Washington Avenue, Albany, NY 12222, USA
FEATURES Location/Qualifiers
source 1..38
```

```

/organism="Erythrocebus patas"
/db_xref="taxon:9538"
/notes="common name:Patas monkey"
join(AF195644.1.<1..274,1..>38)
/product="protamine P2"
join(AF195644.1.1..274,1..38)
/codon_start=1
/product="protamine P2"
/protein_id="AAG42203.1"
/db_xref="GI:11990584"
/translation="MYRYRSLRSPHEVHGQVYQEGHNGQEQGLSPHVEVY
Y"
ERTHQGHSHRRRCSSQRLRIHRRHRRSCRRRRRRRRRRRRGCTRRRCRR
Y"
1..>38
/number=2
15 a 6 c 13 g 4 t
BASE COUNT
ORIGIN

Query Match 68.0%; Score 13.6; DB 91; Length 38;
Best Local Similarity 80.0%; Pred. No. 2.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaaactgcagaagcactga 20
   ||| ||||||| |||
Db 19 GAAGATGCAGAAGTACTAA 38

RESULT 4
SEPROTI12
LOCUS SEPOTI12 38 bp DNA PRI 23-DEC-2000
DEFINITION Semnopithecus entellus protamine 2 gene, exon 2 and complete cds.
ACCESSION AF195643
VERSION AF195643.1 GI:11990579
KEYWORDS
SEGMENT 2 of 2
SOURCE Semnopithecus entellus.
ORGANISM Semnopithecus entellus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Colobinae; Semnopithecus.
REFERENCE 1 (bases 1 to 38)
AUTHORS Karanth,P., Stewart,C.B., Holt,R.A., deKoning,J. and Messier,W.
TITLE Positive Darwinian selection on the lineage leading to humans
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 38)
AUTHORS Karanth,P. and Stewart,C.B.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1999) Biological Sciences, State University of
New York, 1400 Washington Avenue, Albany, NY 12222, USA
FEATURES
source
1..38
/organism="Semnopithecus entellus"
/db_xref="taxon:88029"
/notes="common name:Hanuman langur"
join(AF195642.1.<1..274,1..>38)
/product="protamine 2"
join(AF195642.1.1..274,1..38)
/codon_start=1
/product="protamine 2"
/protein_id="AAG42202.1"
/db_xref="GI:11990580"
/translation="MYRYRMSLSRSPHEVHGQVYQEGHNGQEQGLSPHVEVY
Y"
ERTHGYSHRRRCSSRRRLRIHRRHRRSCRRRRRRRRRRRRGCTRRRCRR
Y"
1..>38
/number=2
15 a 6 c 13 g 4 t
BASE COUNT
ORIGIN

Query Match 68.0%; Score 13.5; DB 97; Length 38;
Best Local Similarity 80.0%; Pred. No. 2.7e+04;

```

```

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaaactgcagaagcactga 20
   ||| ||||||| |||
Db 19 GAAGATGCAGAAGTACTAA 38

RESULT 5
A69037
LOCUS A69037 36 bp DNA PAT 06-MAY-1999
DEFINITION Sequence 25 from Patent WO9803660.
ACCESSION A69037
VERSION A69037.1 GI:4759966
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 36)
AUTHORS Baudu,P., Riviere,M., Audonnet,J. and Bouchardon,A.
TITLE FELINE POLYNUCLEOTIDE VACCINE FORMULA
JOURNAL Patent: WO 9803660-A 25 29-JAN-1998;
BAUDU PHILIPPE (FR)
COMMENT Other publication FR 2751223 19980123.
FEATURES
source
1..36
/organism="unidentified"
/db_xref="taxon:32644"
14 a 6 c 11 g 5 t
BASE COUNT
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 36;
Best Local Similarity 93.3%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 aaactgcagaagcga 16
   ||| ||||||| |||
Db 2 AAATGCAGAAGGAA 16

RESULT 6
AR018055/c
LOCUS AR018055 43 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 12 from patent US 5780286.
ACCESSION AR018055
VERSION AR018055.1 GI:3973658
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 43)
AUTHORS Dillon,P.J. and Vockley,J.G.
TITLE Arginase II
JOURNAL Patent: US 5780286-A 12 14-JUL-1998;
FEATURES
source
1..43
/organism="unknown"
8 a 10 c 8 g 17 t
BASE COUNT
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 43;
Best Local Similarity 93.3%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 actgcagaagcact 18
   ||| ||||||| |||
Db 16 ACTGCAGAAGGCAAT 2

RESULT 7
AR071945/c
LOCUS AR071945 43 bp DNA PAT 18-FEB-2000

```

DEFINITION Sequence 12 from patent US 5912159.
ACCESSION AR071945
VERSION AR071945.1 GI:7222833
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 43)
AUTHORS Vockley,J.G and Dillon,P.J.
TITLE Arginase II
JOURNAL Patent: US 5912159-A 12 15-JUN-1999;
FEATURES Location/Qualifiers
1..43
/organism="unknown"
BASE COUNT 8 a 10 c 8 g 17 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 43;
Best Local Similarity 93.3%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 actgcagaagcact 18
|||||
Db 16 ACTGCAGAAGCAAT 2

RESULT 8
E05905 E05905 44 bp DNA PAT 29-SEP-1997
LOCUS
DEFINITION Primer.
ACCESSION E05905
VERSION E05905.1 GI:2174092
KEYWORDS JP 1993308963-A/3.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 44)
AUTHORS Kimura,S., Ishikawa,H. and Nakamura,H..
TITLE METHOD FOR STABILIZING PROTEIN MOLECULE
JOURNAL Patent: JP 1993308963-A 3 22-NOV-1993;
COMMENT TANPAKU KOGAKU KENKYUSHO:KK
OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1993308963-A/3
PD 22-NOV-1993
PF 08-MAY-1992 JP 1992115877
PI KIMURA SHIGENOBU, ISHIKAWA HIROKI, NAKAMURA HARUKI PC
PC C12N9/22,C07K3/08,C12N15/55,C12P21/00;
CC strandedness: Single;
CC topology: Linear;
FEATURES Location/Qualifiers
1..44
/organism="synthetic construct"
BASE COUNT 17 a 9 c 11 g 7 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 10; Length 44;
Best Local Similarity 93.3%; Pred. No. 3.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 aaactgcagaagcca 16
|||||
Db 4 AAATGCAGACGGCA 18

RESULT 9
I27338 I27338 27 bp DNA PAT 06-FEB-1997
LOCUS
DEFINITION Sequence 13 from patent US 5565320.
ACCESSION I27338

VERSION I27338.1 GI:1818114
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Cimino,G.D. and Lin,L.
TITLE Quality control assay for platelet decontamination
JOURNAL Patent: US 5565320-A 13 15-OCT-1996;
FEATURES Location/Qualifiers
1..27
/organism="unknown"
BASE COUNT 10 a 7 c 6 g 4 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 27;
Best Local Similarity 83.3%; Pred. No. 4.4e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 aactgcagaagcactga 20
|||||
Db 2 AACTGCAGAAGCTAGGA 19

RESULT 10
E04313 E04313 44 bp DNA PAT 29-SEP-1997
LOCUS
DEFINITION PCR primer for introducing point mutation into Escherichia coli
ACCESSION E04313
VERSION E04313.1 GI:2172516
KEYWORDS JP 1993038286-A/3.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 44)
AUTHORS Kimura,S..
TITLE VARIANT TYPE ESCHERICHIA COLI RIBONUCLEASE H
JOURNAL Patent: JP 1993038286-A 3 19-FEB-1993;
COMMENT TANPAKU KOGAKU KENKYUSHO:KK
OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1993038286-A/3
PD 19-FEB-1993
PF 07-AUG-1991 JP 1991197703
PI KIMURA SHIGENOBU
PC C12N9/22,C12N1/21,C12N15/55,(C12N9/22,C12R1:19),(C12N1/21, PC
C12R1:19);
PC (C12N15/55,C12R1:19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No..
FEATURES Location/Qualifiers
1..44
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 15 a 10 c 12 g 7 t
ORIGIN

Query Match 65.0%; Score 13; DB 10; Length 44;
Best Local Similarity 100.0%; Pred. No. 5.4e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 actgcagaagcca 16
|||||
Db 6 ACTGCAGAAGGCA 18

RESULT 11
A45762

LOCUS A45762 22 bp DNA 07-MAR-1997
DEFINITION Sequence 19 from Patent WO9520672.
ACCESSION A45762
VERSION A45762.1 GI:2300149
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 22)
AUTHORS Cerutti,M., Chaabhi,H., Devauchelle,G., Gauthier,L., Kaczorek,M., Lefranc,M. and Poul,M.
TITLE RECOMBINANT BACULOVIRUS AND USE THEREOF IN THE PRODUCTION OF MONOCLONAL ANTIBODIES
JOURNAL Patent: WO 9520672-A 19 03-AUG-1995;
COMMENT Other publication AU 1582495 950815
Other publication FR 2715664 950804.
FEATURES
source 1..22
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 5 a 3 c 7 g 3 t 4 others
ORIGIN

Query Match 64.0%; Score 12.8; DB 9; Length 22;
Best Local Similarity 68.8%; Pred. No. 7.1e+04;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 gaaactgcagaagcca 16
:|||||:|||||:
Db 5 SAAMCTGCAGRAGTCW 20

RESULT 12
AR069418
LOCUS AR069418 22 bp DNA 18-FEB-2000
DEFINITION Sequence 1 from patent US 5891648.
ACCESSION AR069418
VERSION AR069418.1 GI:7220306
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 22)
AUTHORS Martin,M.T., Smith,R.G., Darsley,M.J., Simpson,D.M. and Blackburn,G.F.
TITLE Reaction-based selection for expression of and concentration of catalytic moieties
JOURNAL Patent: US 5891648-A 1 06-APR-1999;
FEATURES
source 1..22
/organism="unknown"
BASE COUNT 7 a 3 c 9 g 3 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 9; Length 22;
Best Local Similarity 87.5%; Pred. No. 7.1e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaaactgcagaagcca 16
|||||:|||||:
Db 5 GAAACTGCAGGAGTCA 20

RESULT 13
I43148
LOCUS I43148 22 bp DNA 07-OCT-1997
DEFINITION Sequence 1 from patent US 5631137.
ACCESSION I43148
VERSION I43148.1 GI:2468392
KEYWORDS

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 22)
AUTHORS Martin,M.T., Smith,R.G., Darsley,M.J., Simpson,D.M. and Blackburn,G.F.
TITLE Reaction-based selection for expression of and concentration of catalytic moieties
JOURNAL Patent: US 5631137-A 1 20-MAY-1997;
FEATURES
source 1..22
/organism="unknown"
BASE COUNT 7 a 3 c 9 g 3 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 10; Length 22;
Best Local Similarity 87.5%; Pred. No. 7.1e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaaactgcagaagcca 16
|||||:|||||:
Db 5 GAAACTGCAGGAGTCA 20

RESULT 14
A08097
LOCUS A08097 28 bp DNA 20-JUL-1993
DEFINITION Synthetic oligonucleotide M2.
ACCESSION A08097
VERSION A08097.1 GI:413338
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 28)
AUTHORS Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE Serine protease inhibitor proteins, medicaments containing them, DNA sequences coding for these proteins and methods for producing these proteins, medicaments and DNA sequences
JOURNAL Patent: EP 0373335-A 19 20-JUN-1990;
FEATURES
source 1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 6 a 8 c 10 g 4 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 9; Length 28;
Best Local Similarity 87.5%; Pred. No. 7.1e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 ctgcagaagcactga 20
|||||:|||||:
Db 7 CAGCAGCAGGACTGA 22

RESULT 15
A08101
LOCUS A08101 28 bp DNA 20-JUL-1993
DEFINITION Synthetic oligonucleotide M6.
ACCESSION A08101
VERSION A08101.1 GI:413342
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 28)
AUTHORS Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
TITLE Serine protease inhibitor proteins, medicaments containing them, DNA sequences coding for these proteins and methods for producing

JOURNAL these proteins, medicaments and DNA sequences
Patent: EP 0373335-A 23 20-JUN-1990;
Gruenenthal GmbH

FEATURES
source Location/Qualifiers
1..28

BASE COUNT 5 a 9 c 9 g 5 t
ORIGIN /organism="synthetic construct"
/db_xref="taxon:32630"

Query Match 64.0%; Score 12.8; DB 9; Length 28;
Best Local Similarity 87.5%; Pred. NO. 7.1e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 ctgcagaaggaactga 20
Db 7 CAGCAGCAGGCACTGA 22

Search completed: October 2, 2001, 15:56:39
Job time: 14163 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:35 ; Search time 876.95 seconds
(without alignments)
14.320 Million cell updates/sec

Title: US-09-757-100B-6

Perfect score: 20
Sequence: 1 gaaactgcagaagcactga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601:*

- 1: /SIDSB8/gcgdata/geneseq/geneseq/NA1980.DAT:*
- 2: /SIDSB8/gcgdata/geneseq/geneseq/NA1981.DAT:*
- 3: /SIDSB8/gcgdata/geneseq/geneseq/NA1982.DAT:*
- 4: /SIDSB8/gcgdata/geneseq/geneseq/NA1983.DAT:*
- 5: /SIDSB8/gcgdata/geneseq/geneseq/NA1984.DAT:*
- 6: /SIDSB8/gcgdata/geneseq/geneseq/NA1985.DAT:*
- 7: /SIDSB8/gcgdata/geneseq/geneseq/NA1986.DAT:*
- 8: /SIDSB8/gcgdata/geneseq/geneseq/NA1987.DAT:*
- 9: /SIDSB8/gcgdata/geneseq/geneseq/NA1988.DAT:*
- 10: /SIDSB8/gcgdata/geneseq/geneseq/NA1989.DAT:*
- 11: /SIDSB8/gcgdata/geneseq/geneseq/NA1990.DAT:*
- 12: /SIDSB8/gcgdata/geneseq/geneseq/NA1991.DAT:*
- 13: /SIDSB8/gcgdata/geneseq/geneseq/NA1992.DAT:*
- 14: /SIDSB8/gcgdata/geneseq/geneseq/NA1993.DAT:*
- 15: /SIDSB8/gcgdata/geneseq/geneseq/NA1994.DAT:*
- 16: /SIDSB8/gcgdata/geneseq/geneseq/NA1995.DAT:*
- 17: /SIDSB8/gcgdata/geneseq/geneseq/NA1996.DAT:*
- 18: /SIDSB8/gcgdata/geneseq/geneseq/NA1997.DAT:*
- 19: /SIDSB8/gcgdata/geneseq/geneseq/NA1998.DAT:*
- 20: /SIDSB8/gcgdata/geneseq/geneseq/NA1999.DAT:*
- 21: /SIDSB8/gcgdata/geneseq/geneseq/NA2000.DAT:*
- 22: /SIDSB8/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	22	Human focal adhesi
2	15	75.0	15	22	Human focal adhesi
3	14.8	74.0	40	21	Dystrophin gene sp
4	14	70.0	33	21	Neisseria species
5	13.8	69.0	25	16	Splice site for mo
6	13.8	69.0	40	17	Soluble mouse Fas
7	13.6	68.0	42	21	HIV Vpr C-terminal
8	13.6	68.0	42	21	HIV-1 fragment PCR
9	13.4	67.0	20	20	PCR primer used to
10	13.4	67.0	36	19	Primer AB067 for F
11	13.4	67.0	43	18	PCR primer 5R used

c 12	13.4	67.0	43	18	AAV00587	PCR primer 5R used
c 13	13.4	67.0	43	19	AAV17634	Homo sapiens argin
c 14	13.4	67.0	43	20	AAV77316	Arginase II cDNA c
c 15	13.4	67.0	43	21	AAAO8070	Human arginase II
c 16	13.4	67.0	44	14	AAQ40649	Ribonuclease prime
c 17	13.4	67.0	44	15	AAQ53496	pJK95N Primer 3.
c 18	13.4	67.0	44	16	AAQ96275	E.coli variant rib
c 19	13.2	66.0	27	14	AAQ47315	PCR primer for mit
c 20	13.2	66.0	27	17	AAAT47908	platelet mitochond
c 21	13.2	66.0	27	21	AAZ58555	HSV-2 antigen UL50
c 22	13	65.0	44	14	AAQ38711	Mutagenic PCR prim
c 23	13	65.0	44	14	AAQ40640	Ribonuclease H mut
c 24	13	65.0	44	20	AAQ34253	phoA-luc construct
c 25	12.8	64.0	21	16	AAQ9581	Primer B7 (Group 4
c 26	12.8	64.0	22	14	AAQ47805	Mouse spleen RNA 5
c 27	12.8	64.0	35	19	AAV49413	Primer AB017 for C
c 28	12.8	64.0	35	22	AAQ90404	Flounder growth ho
c 29	12.8	64.0	36	20	AAQ56735	WO915652 primer 2
c 30	12.8	64.0	36	21	AAQ55237	Rat megasin PCR pri
c 31	12.6	63.0	21	13	AAQ35615	env primer HIV3B10
c 32	12.6	63.0	21	14	AAQ35338	PCR primer HIV3B10
c 33	12.6	63.0	25	20	AAQ10964	PCR primer for HIV
c 34	12.6	63.0	26	16	AAQ11157	HIV gp160 PCR sens
c 35	12.6	63.0	26	18	AAAT8543	gp160 PCR sense pr
c 36	12.6	63.0	26	19	AAV21769	HIV gp160 gene amp
c 37	12.6	63.0	30	20	AAV08912	PCR primer for SIV
c 38	12.6	63.0	36	21	AAZ95630	HIV-1 env gene nat
c 39	12.6	63.0	39	21	AAZ55133	Neisseria species
c 40	12.6	63.0	43	21	AAAT7204	HIV-1 proviral DNA
c 41	12.4	62.0	25	18	AAAT47491	Human Inducible nI
c 42	12.4	62.0	28	21	AAAI4276	PCR primer pGBD-SM
c 43	12.4	62.0	29	19	AAV34689	Yeast YCF1 gene PC
c 44	12.4	62.0	30	21	AAQ30159	PCR primer TL F2 u
c 45	12.4	62.0	35	21	AAZ59858	Human apo C-III pr

ALIGNMENTS

RESULT 1

AAQ65538
ID AAC65538 standard; DNA; 20 BP.

XX AAC65538;

XX 12-FEB-2001 (first entry)

XX Human focal adhesion kinase antisense sequence #4.

XX Human; focal adhesion kinase; FAK; signal transduction; cancer;

XX embryonic development disorder; angiogenic disorder; wound healing;

XX antisense; phosphorothioate; ss.

XX Homo sapiens.

XX US6133031-A.

XX 17-OCT-2000.

XX 19-AUG-1999; 99US-0377310.

XX 19-AUG-1999; 99US-0377310.

XX (ISIS-) ISIS PHARM INC.

XX Monia BP, Gaarde WA;

XX WPI; 2001-006141/01.

XX New antisense compounds for inhibiting focal adhesion kinase

XX expression, especially useful for inhibiting retinal

XX neovascularization, or for diagnosing and treating e.g. colon cancer -

XX

PS Claim 3; Column 23; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, and is implicated
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX Sequence 20 BP; 8 A; 4 C; 6 G; 2 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gaaactgcagaaggcactga 20
 |||||
 Db 1 gaaactgcagaaggcactga 20

RESULT 2

AAC65558

ID AAC65558 standard; DNA; 15 BP.

XX

AC AAC65558;

DT 12-FEB-2001 (first entry)

XX Human focal adhesion kinase antisense sequence #24.

XX Human; focal adhesion kinase; FAK; signal transduction; cancer;

KW embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphorothioate; ss.

XX Homo sapiens.

OS US6133031-A.

PN 17-OCT-2000.

PD 19-AUG-1999; 99US-0377310.

XX 19-AUG-1999; 99US-0377310.

PF (ISIS-) ISIS PHARM INC.

PR Monla BP, Gaarde WA;

XX WPI; 2001-006141/01.

DR New antisense compounds for inhibiting focal adhesion kinase

XX expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

XX Example 2; Column 25; 30pp; English.

PS The present invention describes a number of phosphorothioate antisense

XX sequences to the human focal adhesion kinase (FAK) protein. This protein

CC is involved in integrin-mediated signal transduction, and is implicated

CC in cancer, particularly colon, breast and oral tumours, embryonic

CC development disorders, angiogenic disorders and wound healing. The

CC antisense sequences, including the one shown here, can be used in the

CC treatment of all of these.

Qy 3 aactgcagaaggcac 17
 |||||

Db 1 aactgcagaaggcac 15
 |||||

RESULT 3

AZ48582/C

ID AZ48582 standard; DNA; 40 BP.

XX

AC AZ48582;

DT 31-MAR-2000 (first entry)

XX Dystrophin gene specific primer R5.

DE Muscular dystrophy; rod domain; adeno-associated virus; AAV;

KW dystrophin gene; truncated; PCR primer; ss.

XX Homo sapiens.

OS JP11318467-A.

PN 24-NOV-1999.

PD 08-MAY-1998; 98JP-0142134.

XX 08-MAY-1998; 98JP-0142134.

PF (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

KW (KOKU-) KOKURITSU SEISHIN SHINKAI CENT. SOCHO.

XX WPI; 2000-100771/09.

DR A gene encoding a shortened dystrophin - useful for the treatment of

XX muscular dystrophy

PT Example 1; Page 10; 44pp; Japanese.

XX The invention provides a gene for the treatment of muscular dystrophy

CC having at least one rod repeat structure of hinge 1, hinge 4 and rod

CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene

CC and a gene-introducing medium consisting of an adeno-associated virus

CC (AAV) vector or lentivirus vector containing the rod shortened

CC dystrophin genes can be used for the genetic treatment of muscular

CC dystrophy of low immune reaction. Sequences AAZ48572-584 represent

CC dystrophin gene specific primers.

XX Sequence 40 BP; 4 A; 13 C; 12 G; 11 T; 0 other;

SQ Query Match 74.0%; Score 14.8; DB 21; Length 40;

Best Local Similarity 88.9%; Pred. No. 4.6e+02;

Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 aaactgcagaaggcactg 19
 |||||

Db 21 AGACTGCAGAGGCCCTG 4

RESULT 4

AZ54675

ID AZ54675 standard; DNA; 33 BP.

XX

AC AZ54675;

DT 21-MAR-2000 (first entry)

XX Neisseria species ORF cloning PCR primer #60.

DE Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;

XX antibacterial; gene therapy; PCR primer; ss.

PR 27-MAY-1994; 94US-0250478.

PA (COLS) UNIV COLORADO.

PI Bellgrau D, Duke RC;

XX WPI; 1996-030252/03.

DR Use of Fas ligand - for suppressing lymphocyte-mediated immune responses, e.g. transplant rejection or auto-immune conditions

PS Disclosure; Page 32; 51pp; English.

XX This 5' DNA primer is used for the synthesis of the full-length soluble mouse Fas ligand gene. It is used in conjunction with the 3' DNA primer (AA07687).

XX Sequence 40 BP; 9 A; 13 C; 11 G; 7 T; 0 other;

Query Match 69.0%; Score 13.8; DB 17; Length 40;

Best Local Similarity 88.2%; Pred. No. 1.3e+03;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 3 aactgcagaaggaactg 19

Db 22 acctgcagaaggaactg 38

RESULT 7

AAZ88079/c

ID AAZ88079 standard; DNA; 42 BP.

XX AC AAZ88079;

XX 20-APR-2000 (first entry)

DE HIV Vpr C-terminal and Tat N-terminal PCR primer SEQ ID NO:15.

XX Lentiviral vector; packaging; gag; pol; gene therapy; infection; gene expression; PCR primer; ss.

XX Human immunodeficiency virus type 1.

OS WO200000600-A2.

XX 06-JAN-2000.

XX 26-MAY-1999; 99WO-US11516.

XX 26-MAY-1998; 98US-0086635.

XX (CHAN/) CHANG L.

XX Chang L;

XX WPI; 2000-137067/12.

XX New packaging vector comprising a nucleotide sequence encoding Gag and Pol proteins of a reference lentivirus useful for the delivery of non-lentiviral genes to target cells

XX Example 1; Page 148; 311pp; English.

XX The present invention describes a packaging vector (PV) comprising a nucleotide sequence encoding Gag and Pol proteins of a reference lentivirus that differs from the reference lentivirus at least in that: (a) its major splice donor site is either deleted or is sufficiently different from the reference lentivirus so that it is not a potential site for homologous recombination; and (b) it lacks a functional major packaging signal so that the introduced vector causes the host cell to produce packaging vector particles comprising functional Gag and Pol proteins. The vectors are useful for transforming (eukaryotic) cells to

CC express specific genes at high levels, e.g. for gene therapy. The improved vectors are safer, yet permit increased efficiency of packaging the recombinant viral genome and increased long-term gene expression. These properties are required for gene therapy as a means of treating infectious and non-infectious diseases. Unlike other retroviruses, the lentiviruses are able to infect non-dividing cells. The present sequence represents an HIV Vpr C-terminal and Tat N-terminal PCR primer which is used in the exemplification of the present invention.

XX Sequence 42 BP; 4 A; 15 C; 5 G; 18 T; 0 other;

Query Match 68.0%; Score 13.6; DB 21; Length 42;

Best Local Similarity 80.0%; Pred. No. 1.6e+03;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 gaaactgcagaaggaactg 20

Db 42 GAAAGAGCAGAAGACAGTGA 23

RESULT 8

AAZ56558/c

ID AAZ56558 standard; DNA; 42 BP.

XX AC AAZ56558;

XX 21-MAR-2000 (first entry)

DE HIV-1 fragment PCR primer #2.

XX Lentiviral vector; gene amplification system; gene therapy; packaging; expression; transgenic animal; ss.

XX Human immunodeficiency virus type 1.

OS WO9961598-A2.

XX 02-DEC-1999.

XX 26-MAY-1999; 99WO-US11634.

XX 26-MAY-1998; 98US-0086635.

XX (UVFL) UNIV FLORIDA.

XX Chang L, Cui Y, Iwakuma T;

XX WPI; 2000-072617/06.

XX New gene amplification system, useful to express a target gene in cells e.g. in gene therapy

XX Example 3; Page 98; 197pp; English.

XX The present invention describes a gene amplification system comprising a transducing vector, a packaging vector and a suitable host. The lentiviral vectors are useful to transform suitable host cells to express a target gene at high levels in the cells. For example, the vector system can be used therapeutically with both dividing and non-dividing human cells (e.g. neuronal cells and haematopoietic stem cells) e.g. in gene therapy. The vector system is also useful to produce polypeptides in cells (e.g. TE671 and HeLa cells) in vitro and in vivo, and to produce transgenic animals e.g. animals expressing human proteins. Use of modified lentiviral packaging vectors reduces the risk of generating replication-competent virus through recombination with the transducing vector or a defective provirus endogenous to the host cell, compared with prior art lentivirus packaging vectors. The vector system provides increased long-term gene expression compared with other viral vectors (e.g. could remain transduced for at least 120 days), so it may be possible for the vector to remain transduced for the necessary time course of a particular treatment. AAZ56538 to AAZ54578 represent nucleotide sequences used in the exemplification of

CC the present invention.

XX Sequence 42 BP; 4 A; 15 C; 5 G; 18 T; 0 other;

SQ Query Match 68.0%; Score 13.6; DB 21; Length 42;

Best Local Similarity 80.0%; Pred. No. 1.6e+03;

Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gaaactgcagaagcactga 20

||||| ||||| |||

Db 42 GAAAGAGCAGAAGACACTGA 23

RESULT 9

AAZ04825/C

ID AAZ04825 standard; DNA; 20 BP.

XX AC AAZ04825;

XX 07-OCT-1999 (first entry)

XX PCR primer used to amplify an ORF of Chlamydia trachomatis.

DE Vaccines; eye disease; conventional trachoma; nonendemic trachoma;

KW paratrachoma; inclusion conjunctivitis; genital disease; perihhepatitis;

KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis; PCR primer;

KW bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.

XX Synthetic.

OS Chlamydia trachomatis.

XX WO928475-A2.

XX 10-JUN-1999.

XX 27-NOV-1998; 98WO-1B01939.

XX 04-NOV-1998; 98US-0107077.

XX 28-NOV-1997; 97FR-0015041.

XX 17-DEC-1997; 97FR-0016034.

XX (GEST) GENSET.

XX Griffais R;

XX WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis

XX Disclosure; Page 1720; 1755pp; English.

XX PCR primers AAZ01426-206209 were used to amplify open reading frames

CC (ORFs) of the genome of Chlamydia trachomatis (see AAZ01425). These ORFs

CC encode polypeptides (see AAY36754-Y37949) which can be used as vaccines

CC against Chlamydia trachomatis. Antisense and ribozyme sequences

CC can also be used to control growth of the microorganism. Chlamydia

CC trachomatis is responsible for a large number of diseases, e.g. eye

CC diseases such as conventional trachoma, nonendemic trachoma,

CC paratrachoma, and inclusion conjunctivitis; genital diseases such as

CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,

CC perihhepatitis, bartholinitis; pneumopathy in breast feeding infants;

CC and venereal lymphogranulomatosis. The polypeptides of the

CC invention may be of use in treating these diseases.

XX Sequence 20 BP; 5 A; 5 C; 5 G; 5 T; 0 other;

Query Match

Best Local Similarity 67.0%; Score 13.4; DB 20; Length 20;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 actgcagaagcact 18

Db 19 ACTGCAGAACTCACT 5
||||| |||||

RESULT 10

AAV49285

ID AAV49285 standard; DNA; 36 BP.

XX AC AAV49285;

XX 28-OCT-1998 (first entry)

XX Primer AB067 for FIV env gene.

XX Multivalent vaccine; cat; pathogen; respiratory disease; FeLV; FIV; PCV;

KW digestive disease; feline leukaemia virus; feline panleukopenia virus;

KW feline calicivirus; feline immunodeficiency virus; FIV; rabies virus;

KW vector; envelope glycoprotein; primer; PCR; amplification; ss.

XX Synthetic.

OS Feline immunodeficiency virus.

XX FR2751223-A1.

XX 23-JAN-1998.

XX 19-JUL-1996; 96FR-0009337.

XX 19-JUL-1996; 96FR-0009337.

XX (INMR) RHONE MERIEUX SA.

XX Audonnet JCF, Baudu P, Bouchardon A, Riviere MEA;

XX WPI; 1998-112823/11.

XX Multi-valent polynucleotide vaccines against feline pathogens -

PT consist of at least 3 plasmids able to express protective antigens

PT from specified viruses

XX Example 17; Page 16; 42pp; French.

XX The invention relates to a multivalent vaccine for protecting cats

CC against several pathogens, especially pathogens associated with

CC respiratory and digestive diseases. The pathogens are especially

CC selected from feline leukaemia virus (FeLV), feline panleukopenia

CC virus (FPV), feline calicivirus (FCV), feline immunodeficiency virus

CC (FIV), or rabies virus. The vaccines are preferably composed of

CC polynucleotide sequences encoding 3 antigens, all as part of vectors.

CC Primers AAV49285-V49286 were used to PCR amplify the feline

CC immunodeficiency virus (FIV) Petaluma strain gene encoding the envelope

CC glycoprotein. The sequence sub-cloned into the plasmid pVR1012 to

CC generate plasmid pAB030 for use in the vaccine.

XX Sequence 36 BP; 14 A; 6 C; 11 G; 5 T; 0 other;

Query Match

Best Local Similarity 67.0%; Score 13.4; DB 19; Length 36;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 aaactgcagaagcga 16

||||| ||||| |||

Db 2 aaactgcagaagcga 16

RESULT 11

AAT84971/C

ID AAT84971 standard; DNA; 43 BP.

XX AC AAT84971;

XX 01-APR-1998 (first entry)

XX PCR primer 5R used to amplify arginase II from a cDNA library.
 XX Arginase II; proline production; glutamate production; hyperarginaemia;
 DE nitric oxide biosynthesis; arginase activity; urea cycle disease;
 KW hypertension; hypotension; hyperammonaemia; prostate disease;
 KW PCR primer; ss.
 XX Synthetic.
 OS
 XX WO9733985-A1.
 PN
 XX 18-SEP-1997.
 PD
 XX 14-MAR-1996; 96WO-US03561.
 PF
 XX 14-MAR-1996; 96WO-US03561.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Dillon PJ, Vockley JG;
 XX WPI; 1997-470867/43.
 DR
 XX Polynucleotide encoding human arginase II - useful to treat,
 PT diagnose and monitor, e.g. urea cycle disorders, prostatic disease,
 PT hypotension and nitric oxide mediated immune and nervous diseases
 XX
 PS Example 1; Fig 3; 85pp; English.
 XX
 CC PCR primers AAT84962-74 were used to amplify human arginase II from a
 CC Jurkat cell line cDNA library. The cDNA sequence of Arginase I was used
 CC as a probe sequence for a computer search of cDNA databases. Several
 CC expressed sequence tags with 50-60% sequence homology were identified.
 CC These were combined to give a 1075 bp sequence. The above PCR primers
 CC were designed to the extreme 3' and 5' ends of the consensus sequence,
 CC and used to isolate arginase II. In addition to a hypothetical role in
 CC the production of proline and glutamate, it is postulated that
 CC arginase II may play an important role in nitric oxide biosynthesis
 CC through the production of ornithine as a precursor of glutamate.
 CC Arginase II, or its agonists, antagonists and fragments, are used to
 CC treat conditions associated with lack of arginase activity. Compounds
 CC that inhibit activation of the protein are used to treat conditions
 CC associated with excess arginase activity. Typical conditions that can be
 CC treated are diseases of the urea cycle, hypertension, hypotension
 CC (caused by sepsis or cytokines), episodic hyperammonaemia, defective
 CC synthesis of proline, glutamate, nitric oxide or ornithine,
 CC hyperarginaemia and related spasticity, prostate disease (e.g. cancer,
 CC prostatitis and benign hypertrophy), prostate or kidney damage, also
 CC nitric oxide associated immune and nervous system diseases. The arginase
 CC II cDNA is used to produce recombinant protein and for chromosome
 CC identification, while its fragments are used (as primers and probes) to
 CC detect arginase II-encoding sequences and to diagnose the above
 CC diseases.
 XX
 SQ Sequence 43 BP; 8 A; 10 C; 8 G; 17 T; 0 other;
 Query Match 67.0%; Score 13.4; DB 18; Length 43;
 Best Local Similarity 93.3%; Pred. No. 2e+03; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 actgcagaaggcact 18
 Db 16 ACTGCAGAAGGCAAT 2
 RESULT 12
 AAV00587/C
 ID AAV00587 standard; cDNA: 43 BP.
 XX
 AC AAV00587;

XX 25-MAR-1998 (first entry)
 XX PCR primer 5R used to amplify arginase II from a cDNA library.
 DE Arginase II; proline production; glutamate production; hyperarginaemia;
 KW nitric oxide biosynthesis; arginase activity; urea cycle disease;
 KW hypertension; hypotension; hyperammonaemia; prostate disease;
 KW PCR primer; ss.
 XX Synthetic.
 OS
 XX WO9733986-A1.
 PN
 XX 18-SEP-1997.
 PD
 XX 20-AUG-1996; 96WO-US13455.
 PF
 XX 14-MAR-1996; 96WO-US03561.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Dillon PJ, Vockley JG;
 XX WPI; 1997-470868/43.
 DR
 XX Nucleic acid encoding human arginase II - useful for treating,
 PT diagnosing and monitoring e.g. urea cycle disorders, hypotension,
 PT nitric oxide-mediated immune and nervous diseases, etc
 XX
 PS Example 1; Fig 3; 93pp; English.
 XX
 CC PCR primers AAV00578-90 were used to amplify human arginase II from a
 CC Jurkat cell line cDNA library. The cDNA sequence of Arginase I was used
 CC as a probe sequence for a computer search of cDNA databases. Several
 CC expressed sequence tags with 50-60% sequence homology were identified.
 CC These were combined to give a 1075 bp sequence. The above PCR primers
 CC were designed to the extreme 3' and 5' ends of the consensus sequence,
 CC and used to isolate arginase II. In addition to a hypothetical role in
 CC the production of proline and glutamate, it is postulated that
 CC arginase II may play an important role in nitric oxide biosynthesis
 CC through the production of ornithine as a precursor of glutamate.
 CC Arginase II, or its agonists, antagonists and fragments, are used to
 CC treat conditions associated with lack of arginase activity. Compounds
 CC that inhibit activation of the protein are used to treat conditions
 CC associated with excess arginase activity. Typical conditions that can be
 CC treated are diseases of the urea cycle, hypertension, hypotension
 CC (caused by sepsis or cytokines), episodic hyperammonaemia, defective
 CC synthesis of proline, glutamate, nitric oxide or ornithine,
 CC hyperarginaemia and related spasticity, prostate disease (e.g. cancer,
 CC prostatitis and benign hypertrophy), prostate or kidney damage, also
 CC nitric oxide associated immune and nervous system diseases. The arginase
 CC II cDNA is used to produce recombinant protein and for chromosome
 CC identification, while its fragments are used (as primers and probes) to
 CC detect arginase II-encoding sequences and to diagnose the above
 CC diseases.
 XX
 SQ Sequence 43 BP; 8 A; 10 C; 8 G; 17 T; 0 other;
 Query Match 67.0%; Score 13.4; DB 18; Length 43;
 Best Local Similarity 93.3%; Pred. NO. 2e+03;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 4 actgcagaaggcact 18
 Db 16 ACTGCAGAAGGCAAT 2
 RESULT 13
 AAV17634/C
 ID AAV17634 standard; DNA: 43 BP.

```

XX AAV17634;
XX AC
XX DT
XX DE 03-AUG-1998 (first entry)
XX DE Homo sapiens arginase II gene PCR primer.
XX KW human; arginase II; disease; gene defect; hyperammonaemia;
XX KW diagnosis; susceptibility; urea cycle; hypertension; hypotension;
XX KW episodic; prolina; biosynthesis; defect; glutamate; nitric oxide;
XX KW ornithine; hyperargininaemia; spasticity; growth retardation;
XX KW mental impairment; prostate; cancer; prostatic; benign prostatic;
XX KW hyperplasia; hypertrophy; damage; kidney; PCR primer; ss.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN EP825260-A2.
XX PD
XX PD 25-FEB-1998.
XX PF
XX PF 20-AUG-1997; 97EP-0306360.
XX PR
XX PR 20-AUG-1996; 96US-0700186.
XX PR (HUMA-) HUMAN GENOME SCI INC.
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX PI
XX PI Dillon P, Vockley JG;
XX DR
XX DR WPI; 1998-132255/13.
XX DR
XX DR New human arginase II protein - useful for treating, e.g. diseases
XX PT associated with defect in arginase II gene such as episodic
XX PT hyperammonaemia
XX PS
XX PS Example 1; Fig 3; 46pp; English.
XX CC The sequence is that of a PCR primer (5R) which was synthesised
XX CC based on a potential arginase II EST consensus sequence. It was
XX CC used in the amplification and cloning of the human arginase II gene.
XX CC
XX SQ Sequence 43 BP; 8 A; 10 C; 8 G; 17 T; 0 other;

Query Match 67.0%; Score 13.4; DB 19; Length 43;
Best Local Similarity 93.3%; Pred. No. 2e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 actgcagaagcgact 18
DB 16 ACTGCAGAGGCAAT 2
|||||
|

RESULT 14
AAX77316/c
ID AAX77316 standard; DNA; 43 BP.
XX AC
XX AC AAX77316;
XX DT
XX DT 09-AUG-1999 (first entry)
XX DE
XX DE Arginase II cDNA cloning primer 5R.
XX KW Arginase II; gene therapy; hyperargininemia; immunogen; human;
XX KW urea cycle disorder; hypertension; hypotension; prostate cancer;
XX KW benign prostatic hyperplasia; PCR primer; ss.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN US5912159-A.
XX

```

```

PD 15-JUN-1999.
XX XX
XX PF 20-AUG-1997; 97US-0914981.
XX PR
XX PR 20-AUG-1997; 97US-0914981.
XX PR 20-AUG-1996; 96US-0700186.
XX XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX PI
XX PI Dillon PJ, Vockley JG;
XX DR
XX DR WPI; 1999-357201/30.
XX XX
XX PT Polypeptides with Arginase activity, useful for treating type I
XX PT arginase and disorders resulting in hyperargininemia
XX PS
XX PS Examples; Fig 3; 38pp; English.
XX CC The invention relates to a polypeptide with Arginase II activity. The
XX CC Arginase II polypeptides may be used in gene therapy for the treatment
XX CC of type I arginase disorders resulting in hyperargininemia. The
XX CC polypeptide sequences may also be used as immunogens to produce
XX CC antibodies, these antibodies may then be used to treat numerous diseases
XX CC e.g. urea cycle disorders, hypertension, hypotension, prostate cancer,
XX CC benign prostatic hyperplasia. Sequences AAX77307-319 represent PCR
XX CC primers used for cloning the Arginase II cDNA.
XX SQ Sequence 43 BP; 8 A; 10 C; 8 G; 17 T; 0 other;

Query Match 67.0%; Score 13.4; DB 20; Length 43;
Best Local Similarity 93.3%; Pred. No. 2e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 actgcagaagcgact 18
DB 16 ACTGCAGAGGCAAT 2
|||||
|

RESULT 15
AAA08070/c
ID AAA08070 standard; DNA; 43 BP.
XX AC
XX AC AAA08070;
XX DT
XX DT 21-JUN-2000 (first entry)
XX DE
XX DE Human arginase II PCR primer SEQ ID NO:12.
XX KW Human; arginase II; arginase I; diagnosis; hypotensive; hypertensive;
XX KW uropathic; cytostatic; neuroprotective; gene therapy; hypertension;
XX KW nitric oxide biosynthesis modulator; urea cycle disease; hypotension;
XX KW episodic hyperammonaemia; hyperargininaemia; spasticity; prostatic;
XX KW growth retardation; progressive mental impairment; prostate disease;
XX KW prostate cancer; benign prostatic hyperplasia; hypertrophy;
XX KW prostate damage; kidney disease; kidney damage; PCR primer; ss.
XX OS Homo sapiens.
XX PN
XX PN US6054308-A.
XX PD
XX PD 25-APR-2000.
XX PF
XX PF 15-JUL-1998; 98US-0116115.
XX PR
XX PR 14-MAR-1996; 96US-0013395.
XX PR 20-AUG-1997; 97US-0914981.
XX PR 20-AUG-1996; 96US-0700186.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX

```

PI Dillon PJ, Vockley JG;
 XX WPI; 2000-328355/28.
 XX
 PT Novel human arginase II polypeptides useful for treating urea cycle
 PT diseases, hypertension, hypotension, episodic hyperammonemia, to
 PT control nitric oxide formation and kidney damage -
 XX
 XX
 PS Example 1; Fig 3; 37pp; English.
 XX
 CC The present invention describes human arginase II. Arginase II has
 CC hypotensive, hypertensive, uropathic, cytostatic and neuroprotective
 CC activities, and can be used in gene therapy and as a nitric oxide
 CC biosynthesis modulator. Human arginase II proteins can be used to treat
 CC diseases associated with or caused by a defect in the arginase II gene
 CC or arginase II gene expression, such as, for e.g. urea cycle diseases,
 CC hypertension, hypotension, episodic hyperammonemia, defects in
 CC biosynthesis of proline, glutamate, nitric oxide and ornithine, as well
 CC as hyperargininaemia and its related spasticity, growth retardation,
 CC and progressive mental impairment, and prostate disease, particularly
 CC prostate cancer, prostatitis and benign prostatic hyperplasia or
 CC hypertrophy, and also prostate damage, kidney disease and kidney
 CC damage. It is also used to control nitric oxide formation in an
 CC individual. Arginase II or its fragments, variants or derivatives can be
 CC used as diagnostic reagents for diagnosing arginase II deficiency in an
 CC individual having or suspected of having a defect in the nitric oxide
 CC pathway and the urea cycle. The genes encoding arginase II are used in
 CC gene therapy techniques to treat the above mentioned disorders. It is
 CC also used to deplete systemic arginine levels in an individual. AAA08061
 CC to AAA08073 represent PCR primers used in the cloning of human arginase
 CC II.
 XX
 SQ Sequence 43 BP; 8 A; 10 C; 8 G; 17 T; 0 other;

Query Match 67.0%; Score 13.4; DB 21; Length 43;
 Best Local Similarity 93.3%; Pred. NO. 2e+03;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 4 actgcagaagcact 18
 Db 16 ACTGCAGAGGCAAT 2
 |||||

Search completed: October 2, 2001, 16:18:36
 Job time: 15480 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:00:46 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-6
Perfect score: 20
Sequence: 1 gaaactgcagaagcactga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1CM493 row: a column: 17
 High quality sequence stop: 33.
 Location/Qualifiers

FEATURES

source 1. .33
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3826072"
 /clone_lib="NIH_MGC_58"
 /tissue_type="hyponephroma"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for long-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."
 BASE COUNT 11 a 7 c 7 g 8 t
 ORIGIN

Query Match 63.0%; Score 12.6; DB 139; Length 33;
 Best Local Similarity 78.9%; Pred. No. 8.2e+04;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 aaactgcagaagcactga 20
 ||||| ||||| |||||
 Db 1 AACCTGTAGAAGACCCCTGA 19

RESULT 5

LOCUS N40673 46 bp mRNA EST 22-JAN-1996
 DEFINITION Yw78g11.r1 Soares_Placenta_8to9weeks_2NBHP8to9W Homo sapiens cDNA clone IMAGE:258404 5' similar to gb:J00117 CHORIOGONADOTROPIN BETA CHAIN PRECURSOR (HUMAN);, mRNA sequence.
 ACCESSION N40673
 VERSION N40673.1 GI:1164270
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 46)
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman ,M., Hultman,M., Kucaba,F., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisakis,E., Waterston ,R., Williamson,A., Wohldmann,P. and Wilson,R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence starts: 1
 High quality sequence stops: 1

Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Seq primer: T7
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source 1. .46
 /organism="Homo sapiens"
 /db_xref="GDB:3888014"
 /db_xref="taxon:9606"
 /clone="IMAGE:258404"
 /clone_lib="Soares_Placenta_8to9weeks_2NBHP8to9W"
 /dev_stage="two placentae: one from 8 weeks and another from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: placenta; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGATTGTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Patima Bonaldo."
 BASE COUNT 9 a 9 c 16 g 11 t 1 others
 ORIGIN

Query Match 63.0%; Score 12.6; DB 159; Length 46;
 Best Local Similarity 78.9%; Pred. No. 8.4e+04;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaaactgcagaagcactg 19
 ||||| ||||| |||||
 Db 45 GCAACAGCAGAAGCCCTGTG 27

RESULT 6
 AZ464547/c 47 bp DNA GSS 04-OCT-2000
 LOCUS IM0274A02F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0274A02 F, DNA sequence.
 ACCESSION AZ464547
 VERSION AZ464547.1 GI:10622672
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 REFERENCE 1 (bases 1 to 47)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert length: 10000 Std Error: 0.00
 Plate: 0274 row: A column: 02
 Seq primer: CGTTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 47.
 Location/Qualifiers

FEATURES source 1. .47
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0274A02"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gii47321141gb1AF129072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor-mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 10 a 9 c 7 g 21 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 244; Length 47;
Best Local Similarity 78.9%; Pred. No. 8.4e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 aaactgcagaagcactga 20

Db 25 AAACACTGAAGACATGGA 7

RESULT 7

AA013617/c

LOCUS AA013617 49 bp mRNA EST 21-JAN-1997
DEFINITION mh10d05.r1 Soares mouse placenta 4NDMP13.5 14.5 Mus musculus cDNA clone IMAGE:442089 5' similar to WP:F35G12.9 CE00978 ;, mRNA sequence.

ACCESSION AA013617

VERSION AA013617.1 GI:1474645

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 49)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Willson, R. and Waterston, R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:267425

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

Seq primer: ETPRimer

High quality sequence stop: 1.

Location/Qualifiers

1. .49

/organism="Mus musculus"

/strain="C57BL/6J"

/db.xref="taxon:10090"

/clone="IMAGE:442089"

Query Match 63.0%; Score 12.6; DB 107; Length 50;
Best Local Similarity 78.9%; Pred. No. 8.5e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 aaactgcagaagcactga 20

Db 40 AACCAGCAGAGCCCTGA 22

/clone_lib="Soares mouse placenta 4NDMP13.5 14.5"
/sex="unknown"
/tissue_type="placenta"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: placenta; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 7 a 13 c 15 g 14 t

ORIGIN

Query Match 63.0%; Score 12.6; DB 1; Length 49;
Best Local Similarity 78.9%; Pred. No. 8.5e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gaaactgcagaagcactg 19

Db 22 GACCTGCAGATGGCAGAC 4

RESULT 8

AU102328/c

LOCUS

DEFINITION AU102328 50 bp mRNA EST 05-APR-2001
KAT06002, mRNA sequence.

ACCESSION AU102328

VERSION AU102328.1 GI:13551848

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 50)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo, K., Suyama, A. and Sugano, S.

Fine Structural analysis of transcription start sites of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries

Unpublished (2001)

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S.

Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

Location/Qualifiers

1. .50

/organism="Homo sapiens"

/db.xref="taxon:9606"

/clone="KAT06002"

/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 2 a 18 c 15 g 15 t

ORIGIN

```

RESULT 9
AUI02796
LOCUS      50 bp      mRNA      EST      05-APR-2001
DEFINITION Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
CAS11037, mRNA sequence.
ACCESSION AUI02796
VERSION   AUI02796
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
          H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
          K., Suyama,A. and Sugano,S.
TITLE     Fine Structural analysis of transcription start sites of human
JOURNAL   mRNAs using full-length enriched and 5'-end enriched cDNA libraries
COMMENT   Unpublished (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yszukid@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             source
    source
    1..50
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="CAS11037"
    /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT      19 a      5 c      21 g      5 t

Query Match      62.0%; Score 12.4; DB 107; Length 50;
Best Local Similarity 92.9%; Pred. No. 1.1e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4      actgcagaagggcac 17
      1 |||||
      30 AGTCAGAAGGCAC 43

RESULT 10
AUI02800
LOCUS      50 bp      mRNA      EST      05-APR-2001
DEFINITION Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
CAS11356, mRNA sequence.
ACCESSION AUI02800
VERSION   AUI02800
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
          H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
          K., Suyama,A. and Sugano,S.
TITLE     Fine Structural analysis of transcription start sites of human
JOURNAL   mRNAs using full-length enriched and 5'-end enriched cDNA libraries
COMMENT   Unpublished (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yszukid@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             source
    source
    1..50
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="CAS11037"
    /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT      19 a      5 c      21 g      5 t

Query Match      62.0%; Score 12.4; DB 107; Length 50;
Best Local Similarity 92.9%; Pred. No. 1.1e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4      actgcagaagggcac 17
      1 |||||
      30 AGTCAGAAGGCAC 43

RESULT 10
AUI02800
LOCUS      50 bp      mRNA      EST      05-APR-2001
DEFINITION Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
CAS11356, mRNA sequence.
ACCESSION AUI02800
VERSION   AUI02800
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
          H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
          K., Suyama,A. and Sugano,S.
TITLE     Fine Structural analysis of transcription start sites of human
JOURNAL   mRNAs using full-length enriched and 5'-end enriched cDNA libraries
COMMENT   Unpublished (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yszukid@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             source
    source
    1..50
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="CAS11037"
    /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT      19 a      5 c      21 g      5 t

Query Match      62.0%; Score 12.4; DB 107; Length 50;
Best Local Similarity 92.9%; Pred. No. 1.1e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4      actgcagaagggcac 17
      1 |||||
      30 AGTCAGAAGGCAC 43

```

```

FEATURES             source
    source
    1..50
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="CAS11356"
    /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT      20 a      8 c      16 g      6 t

ORIGIN
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

Query Match      62.0%; Score 12.4; DB 107; Length 50;
Best Local Similarity 92.9%; Pred. No. 1.1e+05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4      actgcagaagggcac 17
      1 |||||
      19 AGTCAGAAGGCAC 32

RESULT 11
AZ779631/c
LOCUS      31 bp      DNA      GSS      16-FEB-2001
DEFINITION 2M0016K13F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGC2M0016K13 F, DNA sequence.
ACCESSION AZ779631
VERSION   AZ779631
KEYWORDS  GSS.
SOURCE    house mouse.
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 31)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
          Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
          M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
          and Wright,D., Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL   plasmid inserts
COMMENT   Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0016 row: K column: 13
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 31.
Location/Qualifiers
    1..31
    /organism="Mus musculus"
    /strain="C57BL/6J"
    /db_xref="taxon:10090"
    /clone="UUGC2M0016K13"
    /clone_lib="Mouse 10kb plasmid UUGCLM library"
    /sex="Male"
    /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
    /note="vector: PWD42nv; Purified genomic DNA from M.
    musculus C57BL/6J (male) was obtained from the Jackson
    Laboratory Mouse DNA Resource
    (http://www.jax.org/resources/documents/dnares/). The DNA
    was hydrodynamically sheared by repeated passage through a
    0.005 inch orifice at constant velocity. The sheared DNA
    was blunt end-repaired with T4 DNA polymerase and T4
    polynucleotide kinase. Adaptor oligonucleotides were
    ligated to the blunt ends in high molar excess. The
    adaptor DNA was purified and size-selected for a 9.5 to
    10.5 kb range using preparative agarose gel

```

electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

7 a 8 c 8 g 8 t

BASE COUNT
ORIGIN

Query Match 61.0%; Score 12.2; DB 249; Length 31;
Best Local Similarity 82.4%; Pred. No. 1.2e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 4 actgcagaagcactga 20
||||| ||||| ||

Db 23 ACTGCTGAAGCAACGA 7

RESULT 12
A2433216
LOCUS 38 bp DNA GSS 03-OCT-2000
DEFINITION 1M0219A01F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0219A01 F, DNA sequence.
ACCESSION A2433216
VERSION A2433216.1 GI:10557229
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 38)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0219 row: A column: 01
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 38.
Location/Qualifiers
1. 38
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0219A01"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

8 a 7 c 6 g 17 t

BASE COUNT
ORIGIN

Query Match 61.0%; Score 12.2; DB 243; Length 38;
Best Local Similarity 82.4%; Pred. No. 1.3e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 3 aactgcagaagcactg 19
||||| ||||| |||||

Db 13 ATCTCTAGAAGCCACTG 29

RESULT 13
AA971669/c
LOCUS 43 bp mRNA EST 07-JUL-1998
DEFINITION Op95c10.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1583634 3' similar to SW:PWP2_HUMAN Q15269 PERIODIC TRYPTOPHAN PROTEIN 2 HOMOLOG.; mRNA sequence.
ACCESSION AA971669
VERSION AA971669.1 GI:3146959
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 43)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 429 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 43
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1583634"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pf7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CGAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."
8 a 8 c 12 g 15 t

FEATURES
source

Query Match 61.0%; Score 12.2; DB 14; Length 43;
Best Local Similarity 82.4%; Pred. No. 1.3e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

BASE COUNT
ORIGIN

QY 2 aaactgcagaagcact 18
 ||| ||||| |||||
 Db 34 AAAATGCAGTACGCAC 18

RESULT 14

AI048043

LOCUS

DEFINITION vn21f12.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA
 clone IMAGE:1021871 5' similar to WP:rl3H5.2 CE03653

RETINAL-BINDING LIKE PROTEIN ; , mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE AUTHORS
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 and Wright,D., Rose,R., Rose,R., Stokes,R., Tingey,A., von Niederhauser,A.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0102 row: P column: 06
 Seq primer: CACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 32.
 Location/Qualifiers
 1..32
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0102P06"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptored DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (g114732114|gbAF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptored mouse DNA was annealed to
 adaptored vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

TITLE

JOURNAL

COMMENT

FEATURES
 source
 1..43
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1021871"
 /clone_lib="Knowles Solter mouse blastocyst B1"
 /tissue_type="blastocyst"
 /dev_stage="embryo (pre-implantation)"
 /lab_host="DH10B"
 /note="Organ: embryo; Vector: pSPORT; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally from mRNA prepared
 from 800 blastocysts. Primer: SalI(dT):
 5'-CGTCCAGCGTCGACCGTTTTTTTTTTT-3'. CDNAs were
 cloned into the NotI/SalI sites of a pSPORT vector (Life
 Technologies). Two different size selections: B1 (larger
 inserts) and B3."

BASE COUNT

ORIGIN

Query Match 61.0%; Score 12.2; DB 15; Length 43;
 Best Local Similarity 82.4%; Pred. No. 1.3e+05;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 actgcagaagcactga 20

||| ||||| |||||

Db 1 AGTGACGAGGCATTGA 17

RESULT 15

AZ826678

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..32

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0102P06"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptored DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (g114732114|gbAF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptored mouse DNA was annealed to
 adaptored vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

QY

Db

Search completed: October 2, 2001, 15:00:48
Job time: 10817 sec

.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:39 ; Search time 3399.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-7

Perfect score: 20

Sequence: 1 ttctcccttcggtattatt 20

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba1.*

2: gb_ba2.*

3: gb_ba3.*

4: gb_in1.*

5: gb_in2.*

6: gb_in3.*

7: gb_om.*

8: gb_ov.*

9: gb_pat1.*

10: gb_pat2.*

11: gb_ph.*

12: gb_pl1.*

13: gb_pl2.*

14: gb_pl3.*

15: gb_pl4.*

16: em_ba1.*

17: em_ba2.*

18: em_fun.*

19: em_htgo_hum.*

20: em_htgo_inv.*

21: em_htgo_rod.*

22: em_htg_hum1.*

23: em_htg_hum2.*

24: em_htg_hum3.*

25: em_htg_hum4.*

26: em_htg_hum5.*

27: em_htg_hum6.*

28: em_htg_hum7.*

29: em_htg_hum8.*

30: em_htg_inv1.*

31: em_htg_inv2.*

32: em_htg_other.*

33: em_htg_rod.*

34: em_hum1.*

35: em_hum2.*

36: em_hum3.*

37: em_hum4.*

38: em_hum5.*

39: em_hum6.*

40: em_hum7.*

41: em_in.*

42: em_om.*

43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vi2.*
60: gb_htgi.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_ro1.*
95: gb_ro2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	13.8	69.0	33	9 AX060862	Sequence
2	13.8	69.0	41	9 AR109090	Sequence
3	12.6	63.0	20	9 A81012	Sequence
4	12.6	63.0	20	9 A95391	Sequence
5	12.6	63.0	30	9 AR004745	Sequence
6	12.6	63.0	30	9 AR008231	Sequence
7	12.6	63.0	30	10 I77015	Sequence
8	12.6	63.0	30	10 I81010	Sequence

```

9      12.6 63.0 30 10 I81106
10     12.2 61.0 22 20 AR110286
11     12.2 61.0 22 10 I73726
12     12.2 61.0 35 10 I03808
13     12.2 60.0 26 9 AX043897
14     12.2 60.0 26 9 AX059414
15     12.2 60.0 27 9 AX043901
16     12.2 60.0 30 9 A83876
17     12.2 60.0 30 9 AR061353
18     12.2 60.0 30 9 AR108252
19     12.2 60.0 30 10 I16209
20     12.2 60.0 30 10 I66695
21     12.2 60.0 30 10 I84789
22     12.2 60.0 31 9 A04979
23     12.2 60.0 31 9 A04980
24     12.2 60.0 34 9 AR036378
25     12.2 60.0 34 10 I72126
26     12.2 60.0 39 10 AX080565
27     12.2 60.0 40 10 I15938
28     12.2 60.0 40 10 I96137
29     12.2 60.0 42 9 AR104422
30     12.2 60.0 43 9 A46553
31     12.2 60.0 45 9 A18117
32     12.2 60.0 47 9 A82690
33     12.2 60.0 47 9 A82705
34     12.2 60.0 50 10 I34292
35     11.8 59.0 23 9 A81700
36     11.8 59.0 24 10 I51986
37     11.8 59.0 27 9 AR039578
38     11.8 59.0 45 9 AR089681
39     11.6 58.0 18 9 AR030029
40     11.6 58.0 19 9 A64833
41     11.6 58.0 21 10 AX095134
42     11.6 58.0 22 9 AR022302
43     11.6 58.0 22 9 AR066394
44     11.6 58.0 22 9 AR084122
45     11.6 58.0 22 9 AX011523

```

ALIGNMENTS

```

RESULT 1
AX060862
LOCUS AX060862 33 bp DNA
DEFINITION Sequence 83 from Patent WO0078815.
ACCESSION AX060862
VERSION AX060862.1 GI:12406240
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 33).
AUTHORS Huse,W.D. and Wu,H.
TITLE Anti- $\gamma_3(a)$  v $\gamma_3(b)$ ? recombinant human antibodies, nucleic acids
JOURNAL encoding same and methods of use
PATENT: WO 0078815-A 83 28-DEC-2000;
APPLIED Molecular Evolution (US)
FEATURES
Source
1..33
/organism="synthetic construct"
/db_xref="taxon:32630"
<1..>33
/note="Mutated complementarity determining region (CDR)"
CDS
/codon_start=1
/transl_table=11
/protein_id="CAC24917.1"
/db_xref="GI:12406241"
/translation="ILLIRYSQSSIS"
BASE COUNT 4 a 13 c 2 g 14 t
ORIGIN

```

```

Query Match 69.0%; Score 13.8; DB 9; Length 33;
Best Local Similarity 88.2%; Pred. No. 1.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 ttcctccgttattctt 20
   ||| ||||| |||||
DB 3 TCTCATCCGTTATCTT 19

RESULT 2
AR109090 ARI09090 41 bp DNA PAT 14-FEB-2001
LOCUS ARI09090 Sequence 49 from patent US 6114116.
DEFINITION
ACCESSION ARI09090
VERSION ARI09090.1 GI:12825366
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Lemieux,B., Landry,B.S. and Sapolsky,R.J.
TITLE Brassica polymorphisms
JOURNAL Patent: US 6114116-A 49 05-SEP-2000;
FEATURES
Location/Qualifiers
source
1..41
/organism="unknown"
BASE COUNT 8 a 11 c 3 g 18 t 1 others
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 41;
Best Local Similarity 78.9%; Pred. No. 1.2e+04;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctccctccgttattct 19
   ||||| ||| |||||
DB 12 TTTCAGTCTTATCT 30

RESULT 3
AR1012
LOCUS AR1012 20 bp DNA PAT 21-JAN-2000
DEFINITION Sequence 64 from Patent EP0918091.
ACCESSION AR1012
VERSION AR1012.1 GI:6731585
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 20)
AUTHORS Kahn,A. and Chelly,J.
TITLE A gene called XLIS and the XLIS gene product, called doublecortin
JOURNAL and their applications
PATENT: EP 0918091-A 64 26-MAY-1999;
INST NAT SANTE RECH MED (FR)
FEATURES
Location/Qualifiers
source
1..20
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1 a 7 c 0 g 12 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 20;
Best Local Similarity 78.9%; Pred. No. 4.9e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttctccctccgttattct 19
   ||| ||||| |||
DB 2 TTATCCCTTCCTTCTCT 20

```

RESULT 4
LOCUS A95391 20 bp DNA PAT 26-JAN-2000
DEFINITION Sequence 64 from Patent WO9927089.
ACCESSION A95391
VERSION A95391.1 GI:6779435
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 20)
AUTHORS Francis, F. and Kahn, A.
TITLE A GENE CALLED XLIS AND THE XLIS GENE PRODUCT, CALLED DOUBLECORTIN AND THEIR PREPARATIONS
JOURNAL Patent: WO 9927089-A 64 03-JUN-1999;
INST NAT SANTE RECH MED (FR); FRANCIS FIONA (FR)
FEATURES
source Location/Qualifiers
1..20
/organism="Homo sapiens"
/db_xref="taxon:9606"
1 a 7 c 12 t
BASE COUNT 1 a 7 c 12 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 20;
Best Local Similarity 78.9%; Pred. No. 4.9e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttctcccttcgcgttattctt 19
||||| ||||| || ||||
Db 2 TTATCCCTTCCTCTCTCT 20

RESULT 5
LOCUS AR004745 30 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 75 from patent US 5747282.
ACCESSION AR004745
VERSION AR004745.1 GI:3965624
KEYWORDS human.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Skolnick, M.H., Goldgar, D.E., Miki, Y., Swenson, J., Kamb, A., Harshman, K.D., Shattuck-Eidens, D.M., Tavtigian, S.V., Wiseman, R.W. and Futreal, P. Andrew.
TITLE 170-linked breast and ovarian cancer susceptibility gene
JOURNAL Patent: US 5747282-A 75 05-MAY-1998;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
1 a 14 c 2 g 13 t
BASE COUNT 1 a 14 c 2 g 13 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 30;
Best Local Similarity 78.9%; Pred. No. 5e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ttctcccttcgcgttattctt 20
||||| ||||| || |||||
Db 8 TCTCTCTCTCTCTCTCTT 26

RESULT 6
LOCUS AR008231 30 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 75 from patent US 5753441.
ACCESSION AR008231
VERSION AR008231.1 GI:3967340
KEYWORDS human.

SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Skolnick, M.H., Goldgar, D.E., Miki, Y., Swenson, J., Kamb, A., Harshman, K.D., Shattuck-Eidens, D.M., Tavtigian, S.V., Wiseman, R.W. and Futreal, P. Andrew.
TITLE 170-linked breast and ovarian cancer susceptibility gene
JOURNAL Patent: US 5753441-A 75 19-MAY-1998;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
1 a 14 c 2 g 13 t
BASE COUNT 1 a 14 c 2 g 13 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 30;
Best Local Similarity 78.9%; Pred. No. 5e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ttctcccttcgcgttattctt 20
||||| ||||| || |||||
Db 8 TCTCTCTCTCTCTCTT 26

RESULT 7
LOCUS I77015 30 bp DNA PAT 03-APR-1998
DEFINITION Sequence 75 from patent US 5693473.
ACCESSION I77015
VERSION I77015.1 GI:3013169
KEYWORDS human.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Shattuck-Eidens, D.M., Simard, J., Durocher, F., Emi, M. and Nakamura, Y.
TITLE Linked breast and ovarian cancer susceptibility gene
JOURNAL Patent: US 5693473-A 75 02-DEC-1997;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
1 a 14 c 2 g 13 t
BASE COUNT 1 a 14 c 2 g 13 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 10; Length 30;
Best Local Similarity 78.9%; Pred. No. 5e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ttctcccttcgcgttattctt 20
||||| ||||| || |||||
Db 8 TCTCTCTCTCTCTCTT 26

RESULT 8
LOCUS I81010 30 bp DNA PAT 10-JUN-1998
DEFINITION Sequence 75 from patent US 5709999.
ACCESSION I81010
VERSION I81010.1 GI:3209300
KEYWORDS human.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Shattuck-Eidens, D.M., Simard, J., Durocher, F., Emi, M. and Nakamura, Y.
TITLE Linked breast and ovarian cancer susceptibility gene
JOURNAL Patent: US 5709999-A 75 20-JAN-1998;
FEATURES Location/Qualifiers
source 1..30

```

BASE COUNT      1 a      14 c      2 g      13 t
ORIGIN

Query Match      63.0%; Score 12.6; DB 10; Length 30;
Best Local Similarity 78.9%; Pred. No. 5e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tctccctccggtattctt 20
    |||| ||||| | |||||
Db 8 TCTCTCTCTCTCTCTTCTT 26

RESULT 9
LOCUS      I81106      30 bp      DNA      PAT      10-JUN-1998
DEFINITION Sequence 75 from patent US 5710001.
ACCESSION  I81106
VERSION     I81106.1 GI:3209396
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 30)
AUTHORS   Skolnick,M.H., Goldgar,D.E., Miki,Y., Swenson,J., Kamb,A.,
           Harshman,K.D., Shattuck-Eidens,D.M., Tavtigian,S.V., Wiseman,R.W.
           and Futreal,P.Andrew.
TITLE     17q-linked breast and ovarian cancer susceptibility gene
JOURNAL   Patent: US 5710001-A 75 20-JAN-1998;
FEATURES   Location/Qualifiers
            1..30
            /organism="unknown"

BASE COUNT      1 a      14 c      2 g      13 t
ORIGIN

Query Match      63.0%; Score 12.6; DB 10; Length 30;
Best Local Similarity 78.9%; Pred. No. 5e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tctccctccggtattctt 20
    |||| ||||| | |||||
Db 8 TCTCTCTCTCTCTCTTCTT 26

RESULT 10
LOCUS      AR110286      22 bp      DNA      PAT      14-FEB-2001
DEFINITION Sequence 38 from patent US 6114502.
ACCESSION  AR110286
VERSION     AR110286.1 GI:12826562
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 22)
AUTHORS   North,M., Nishina,P., Naggert,J. and Noben-Trauth,K.
TITLE     Gene family associated with neurosensory defects
JOURNAL   Patent: US 6114502-A 38 05-SEP-2000;
FEATURES   Location/Qualifiers
            1..22
            /organism="unknown"

BASE COUNT      2 a      10 c      1 g      9 t
ORIGIN

Query Match      61.0%; Score 12.2; DB 9; Length 22;
Best Local Similarity 82.4%; Pred. No. 7.8e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 tctccctccggtatttc 18
    |||| ||||| | |||

```

```

Db 2 TCTCCCTTCTCTTCTTC 18

RESULT 11
LOCUS      I73726      22 bp      DNA      PAT      03-APR-1998
DEFINITION Sequence 4 from patent US 5686598.
ACCESSION  I73726
VERSION     I73726.1 GI:3009867
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 22)
AUTHORS   North,M., Nishina,P. and Naggert,J.
TITLE     Genes associated with retinal dystrophies
JOURNAL   Patent: US 5686598-A 4 11-NOV-1997;
FEATURES   Location/Qualifiers
            1..22
            /organism="unknown"

BASE COUNT      2 a      10 c      1 g      9 t
ORIGIN

Query Match      61.0%; Score 12.2; DB 10; Length 22;
Best Local Similarity 82.4%; Pred. No. 7.8e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 tctccctccggtatttc 18
    |||| ||||| | |||
Db 2 TCTCCCTTCTCTTCTTC 18

RESULT 12
LOCUS      I03808      35 bp      PAT      02-DEC-1994
DEFINITION Sequence 11 from Patent EP 0055942.
ACCESSION  I03808
VERSION     I03808.1 GI:592013
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 35)
AUTHORS   Inouye,M. and Nakamura,K.
TITLE     Plasmid cloning vehicles
JOURNAL   Patent: EP 0055942-A2 11 14-JUL-1982;
FEATURES   Location/Qualifiers
            1..35
            /organism="unknown"

BASE COUNT      14 a      2 c      10 g      9 t
ORIGIN

Query Match      61.0%; Score 12.2; DB 10; Length 35;
Best Local Similarity 82.4%; Pred. No. 7.9e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttctccctccggtattt 17
    ||| ||||| | |||||
Db 32 TTCCCCCTTTCATTATT 16

RESULT 13
LOCUS      AX043897/c      26 bp      DNA      PAT      23-NOV-2000
DEFINITION Sequence 10 from Patent WO0049166.
ACCESSION  AX043897
VERSION     AX043897.1 GI:11342482
KEYWORDS   .
SOURCE     synthetic construct.
ORGANISM   synthetic construct.
           artificial sequence.

```

REFERENCE 1 (bases 1 to 26)
AUTHORS Graham,F.L., Ng,P., Parks,R., Bacchetti,S. and Anglana,M.
TITLE A system for production of helper dependent adenovirus vectors based on use of endonucleases
JOURNAL Patent: WO 0049166-A 10 24-AUG-2000;
Merck & Co., Inc. (US)

FEATURES
source
1. .26
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 17 a 5 c 3 g 1 t

ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 26;
Best Local Similarity 75.0%; Pred. No. 9.9e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
||||| ||| ||||| ||| |||
Db 26 TTCTACTTTTCGTTTTCCTT 7

RESULT 14
AX059414/c
LOCUS AX059414 26 bp DNA PAT 17-JAN-2001
DEFINITION Sequence 147 from Patent WO0055325.
ACCESSION AX059414
VERSION AX059414.1 GI:12311519
KEYWORDS
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 11;
Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 26)
AUTHORS Preuss,D., Copenhaver,G. and Keith,K.
TITLE Plant chromosome compositions and methods
JOURNAL Patent: WO 0055325-A 147 21-SEP-2000;
The University of Chicago (US)

FEATURES
source
1. .26
Location/Qualifiers
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

BASE COUNT 13 a 1 c 10 g 2 t

ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 26;
Best Local Similarity 75.0%; Pred. No. 9.9e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
||||| ||| ||||| ||| |||
Db 20 TTCTTCCTCTCTCTCTCTCTT 1

RESULT 15
AX043901/c
LOCUS AX043901 27 bp DNA PAT 23-NOV-2000
DEFINITION Sequence 14 from Patent WO0049166.
ACCESSION AX043901
VERSION AX043901.1 GI:11342484
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.

REFERENCE 1 (bases 1 to 27)
AUTHORS Graham,F.L., Ng,P., Parks,R., Bacchetti,S. and Anglana,M.
TITLE A system for production of helper dependent adenovirus vectors based on use of endonucleases
JOURNAL Patent: WO 0049166-A 14 24-AUG-2000;

FEATURES
source
1. .27
Location/Qualifiers
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="cDNA"

BASE COUNT 18 a 5 c 3 g 1 t

ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 27;
Best Local Similarity 75.0%; Pred. No. 9.9e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
||||| ||| ||||| ||| |||
Db 27 TTCTACTTTTCGTTTTCCTT 8

Search completed: October 2, 2001, 15:56:41
Job time: 14165 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:00:48 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-7

Perfect score: 20

Sequence: 1 ttctccctccgtattctt 20

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	14.2	71.0	35	258	TA124F02Q	AL463359 T. brucei
C 2	13.6	68.0	49	250	AZ810736	2M0076B05
C 3	12.8	64.0	37	8	AA500031	vi96e11.r
C 4	12.6	63.0	30	244	A2447206	LM0244407
C 5	12.6	63.0	50	249	AZ776590	2M0010A23
C 6	12.2	61.0	37	243	AZ404740	LM00173D10
C 7	12.2	61.0	39	242	AZ357911	LM00099K18
C 8	12.2	61.0	46	189	T49694	ya78g11.s1
C 9	12.2	61.0	49	24	A1747625	u119d03.x
C 10	12.2	60.0	29	242	AZ385263	LM0143109
C 11	12.2	60.0	31	19	A1355496	qul5c07.x
C 12	12.2	60.0	32	250	AZ850974	2M0153109
C 13	12.2	60.0	34	21	A1551509	vo54h02.x
C 14	12.2	60.0	34	242	AZ357613	LM0099K18
C 15	12.2	60.0	35	250	AZ829796	2M0107H11
C 16	12.2	60.0	37	16	A1125022	ao10c10.s
C 17	12.2	60.0	43	244	AZ483277	LM0308P15
C 18	12.2	60.0	46	20	A1419045	tf53c09.x
C 19	12.2	60.0	47	244	AZ486785	IM0315C23
C 20	12.2	60.0	50	107	AU107386	AU107386
C 21	12.2	60.0	50	249	AZ785697	2M0029A20
C 22	11.8	59.0	27	244	AZ482018	LM0306116
C 23	11.8	59.0	40	11	AA739457	vv54a03.r
C 24	11.8	59.0	43	12	AA815634	vt03d02.r
C 25	11.6	58.0	22	249	AZ806493	2M0068M06
C 26	11.6	58.0	29	244	AZ467429	2M0278M13
C 27	11.6	58.0	29	245	AZ508227	LM0350A24
C 28	11.6	58.0	34	258	TA223C10P	AL480256 T. brucei
C 29	11.6	58.0	37	258	HSMC09A10	X88437 H. sapiens B
C 30	11.6	58.0	39	249	AZ795977	AZ795977 2M0051C03
C 31	11.6	58.0	40	22	A1583566	ts15g06.x
C 32	11.6	58.0	46	12	AA828900	od75h11.s
C 33	11.6	58.0	47	189	T61777	vb93g06.r1
C 34	11.6	58.0	47	249	AZ793185	2M0046P12
C 35	11.6	58.0	48	113	AA248284	2819814.5
C 36	11.6	58.0	48	244	AZ477776	LM0297L24
C 37	11.4	57.0	45	250	AZ837415	2M0132K08
C 38	11.2	56.0	25	21	A1539187	lp78h07.x
C 39	11.2	56.0	26	250	AZ809250	2M0073J08
C 40	11.2	56.0	28	24	A1748505	sb53h08.y
C 41	11.2	56.0	34	16	A1130467	SWOVU3CAN
C 42	11.2	56.0	38	258	TA72601Q	AL457973 T. brucei
C 43	11.2	56.0	40	4	AA226145	nc09f09.r
C 44	11.2	56.0	43	20	A1442931	sa29a07.x
C 45	11.2	56.0	45	244	AZ452742	LM0252G17

ALIGNMENTS

RESULT 1	TA124F02Q/c	TA124F02Q	35 bp	DNA	GSS	13-DEC-2000
LOCUS	T. brucei sheared genomic DNA clone 124f02, reverse sequence, genomic survey sequence.					
DEFINITION	AL463359					
ACCESSION	AL463359.1	GI:11833869				
VERSION	GSS.					
KEYWORDS	Trypanosoma brucei.					
SOURCE	Trypanosoma brucei					
ORGANISM	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;					
REFERENCE	1 (bases 1 to 35)					
AUTHORS	Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.					
TITLE	Direct Submission					

JOURNAL

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk

COMMENT

Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREG927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

FEATURES

Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

source

Location/Qualifiers
1. 35
/organism="Trypanosoma brucei"
/strain="TREG927"
/db_xref="taxon:5691"
/clone="124f02"

BASE COUNT 14 a 2 c 11 g 8 t
ORIGIN

Query Match

71.0%; Score 14.2; DB 258; Length 35;

Best Local Similarity

84.2%; Pred. No. 7.1e+03;

Matches

16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY

2 tctccctccgttattctt 20

Db

28 TCTCCCTCCGTATTATT 10

RESULT 2

AZ810736

LOCUS

AZ810736 49 bp DNA GSS 20-FEB-2001

DEFINITION

2M0076B05R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0076B05 R, DNA sequence.

ACCESSION

AZ810736

VERSION

AZ810736.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 49)

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

JOURNAL

University of Utah Genome Center

COMMENT

University of Utah

JOURNAL

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

COMMENT

Tel: 801 585 5606

JOURNAL

Fax: 801 585 7177

COMMENT

Email: ddunn@genetics.utah.edu

JOURNAL

Insert length: 10000

COMMENT

Std Error: 0.00

JOURNAL

Plate: 0076

COMMENT

row: B

JOURNAL

column: 05

COMMENT

Seq primer: CACACAGGAACAGCTATGACC

JOURNAL

Class: plasmid ends

COMMENT

High quality sequence stop: 49.

JOURNAL

Location/Qualifiers

source

1. 49

JOURNAL

/organism="Mus musculus"

COMMENT

/strain="C57BL/6J"

JOURNAL

/db_xref="taxon:10090"

```

/clone="UUGC2M0076B05"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      4 a      13 c      0 g      32 t
ORIGIN

Query Match      68.0%; Score 13.6; DB 250; Length 49;
Best Local Similarity 80.0%; Pred. No. 1.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
    |||| |||| |||| ||||
Db 9 TTCTTCTTCATTATTCTT 28

RESULT 3
AA500031
LOCUS      AA500031      37 bp      mRNA      EST      01-JUL-1997
DEFINITION clone IMAGE:920108 5' similar to SW:COX1_RAT P05503 CYTOCHROME C
            OXIDASE POLYPEPTIDE I ; mRNA sequence.
ACCESSION  AA500031
VERSION     AA500031.1 GI:2234998
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE   1 (bases 1 to 37)
AUTHORS     Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE       The WashU-HMI Mouse EST Project
JOURNAL     Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:532324
Trace considered overall poor quality
possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES
            Location/Qualifiers
                source
                    1..37

```

```

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:920108"
/clone_lib="Barstead mouse pooled organs MPLRB4"
/sex="mixed"
/tissue_type="pooled organs"
/dev_stage="7 day"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGAATCTGAAGTGGAGCGCCGCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [GTTGGATTCGGTACC], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead."
BASE COUNT      6 a      8 c      5 g      18 t
ORIGIN

Query Match      64.0%; Score 12.8; DB 8; Length 37;
Best Local Similarity 87.5%; Pred. No. 3.3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtatt 16
    |||| |||| ||||
Db 2 TTCTCCCTTCATTAT 17

RESULT 4
AA447206
LOCUS      AA447206      30 bp      DNA      GSS      04-OCT-2000
DEFINITION lM0244J07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0244J07 F, DNA sequence.
ACCESSION  AA447206
VERSION     AA447206.1 GI:10598962
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 30)
AUTHORS     Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE       Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL     Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0244 row: J column: 07
            Seq primer: CGTTGTAAACGACGCCAGT
            Class: plasmid ends
            High quality sequence stop: 30.
FEATURES
            Location/Qualifiers
                source
                    1..30
                    /organism="Mus musculus"
                    /strain="C57BL/6J"
                    /db_xref="taxon:10090"
                    /clone="UUGC1M0244J07"
                    /clone_lib="Mouse 10kb plasmid UUGC1M library"
                    /sex="Male"
                    /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                    /note="vector: PWD42nv; Purified genomic DNA from M.

```

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMW42 (gll14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 2 a 16 c 0 g 12 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 244; Length 30;
Best Local Similarity 78.9%; Pred. No. 3.9e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattctt 19
||||| ||||| | |||||
Db 10 TTCTCTCTCTCTCTCTCT 28

RESULT 5

AZ776590 50 bp DNA GSS 16-FEB-2001
LOCUS
DEFINITION 2M0010A23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0010A23 F, DNA sequence.

ACCESSION AZ776590
VERSION
KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 50)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhauser,A.
and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL

COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0010 row: A column: 23

Seq primer: CGTTGTAACACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 50.

Location/Qualifiers

FEATURES

source

1. .50
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0010A23"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMW42 (gll14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 3 a 22 c 1 g 24 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 249; Length 50;
Best Local Similarity 78.9%; Pred. No. 4.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 tctcccttcggtattctt 20
||||| ||||| | |||||
Db 27 TTCTCTACCTTTTCTTCT 45

RESULT 6

AZ404740 37 bp DNA GSS 03-OCT-2000
LOCUS
DEFINITION 1M0173D10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0173D10 F, DNA sequence.

ACCESSION AZ404740
VERSION
KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 37)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhauser,A.
and Wright,D.,Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL

COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0173 row: D column: 10

Seq primer: CGTTGTAACACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 37.

Location/Qualifiers

FEATURES

source

1. .37
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0173D10"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 5 a 8 c 5 g 19 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 243; Length 37;
 Best Local Similarity 82.4%; Pred. No. 6.3e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttctcccttcgatt 17
 Db 20 TTTCTTTTTCGGTGATT 36

RESULT 7
 LOCUS AZ357911/c 39 bp DNA GSS 02-OCT-2000
 DEFINITION IM0099L13R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 clone UUGCLM0099L13 R, DNA sequence.
 ACCESSION AZ357911
 VERSION AZ357911.1 GI:10471611
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 39)
 REFERENCE Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Kelly
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0099 row: L column: 13
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 39.
 Location/Qualifiers
 1. 39
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM0099L13"
 /clone_1lb="Mouse 10kb plasmid UUGCLM library"

/sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 17 a 0 c 16 g 6 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 242; Length 39;
 Best Local Similarity 82.4%; Pred. No. 6.4e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttctcccttcgatt 17
 Db 17 TTTCTTTTTCGGTGATT 1

RESULT 8
 LOCUS T49694 46 bp mRNA EST 06-FEB-1995
 DEFINITION ya78g11.s1 Stratagene placenta (#937225) Homo sapiens cDNA clone
 IMAGE:67844 3' similar to gb:M24194 GUANINE
 NUCLEOTIDE-BINDING PROTEIN BETA SUBUNIT-LIKE PROTEIN (HUMAN), mRNA
 sequence.
 ACCESSION T49694
 VERSION T49694.1 GI:651554
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 46)
 REFERENCE Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
 Chissoe, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins
 M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore
 B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
 Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
 Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
 Generation and analysis of 280,000 human expressed sequence tags
 Genome Res. 6 (9), 807-828 (1996)
 97044478
 Other ESTs: ya78g11.r1
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 711
 High quality sequence starts: 1 High quality sequence stops: 1
 Source: IMAGE Consortium, LLNL This clone is available royalty-free
 through LLNL; contact the IMAGE Consortium (info@image.llnl.gov)
 for further information. Trace considered overall poor quality
 Insert Length: 711 Std Error: 0.00
 Seq primer: -21ml3
 High quality sequence stop: 1.

FEATURES

Location/Qualifiers
1. .46

/organism="Homo sapiens"
/db_xref="GDB:489509"
/db_xref="taxon:9606"
/clone="IMAGE:67844"
/clone_lib="Stratagene placenta (#937225)"
/sex="male"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: placenta; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Caucasian. Average insert size: 1.2 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' "

BASE COUNT 5 a 13 c 6 g 18 t 4 others
ORIGIN

Query Match 61.0%; Score 12.2; DB 189; Length 46;
Best Local Similarity 73.7%; Pred. No. 6.5e+04;
Matches 14; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 19
||||| ||||| ||| |||
Db 26 TTCTGCTTCNGTTCNCT 44

RESULT 9

AI747625 49 bp mRNA EST 22-JUN-1999
LOCUS ul19d03.x1 sugano mouse embryo mewa Mus musculus cDNA clone
DEFINITION IMAGE:2088005 3' similar to TR:089050 O89050 MUSKELIN. ; , mRNA
sequence.

ACCESSION AI747625
VERSION AI747625.1 GI:5125827
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 49)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:995689

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: custom primer used
High quality sequence stop: 1.

FEATURES

Location/Qualifiers
1. .49

/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2088005"
/clone_lib="Sugano mouse embryo mewa"
/dev_stage="embryo, 14 dpc"
/lab_host="DH10B"
/note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCCTTTTTTTTTTTTTTTT];

double-stranded cDNA was ligated to a DraIII adaptor [TCTTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science).

Custom primers for sequencing: 5' end primer CTCTCTCTCTAAAGCTGGG and 3' end primer CGACCTCAGCTGACGACA."

BASE COUNT 11 a 11 c 5 g 22 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 24; Length 49;
Best Local Similarity 82.4%; Pred. No. 6.6e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 tcctcttcggtattctt 20
||||| ||||| ||| |||
Db 28 TCCTCTCTCTTTTCTT 44

RESULT 10

AZ385263/c 29 bp DNA GSS 02-OCT-2000
LOCUS IM0143109R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0143109 R, DNA sequence.

ACCESSION AZ385263
VERSION AZ385263.1 GI:10498963
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 29)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss University of Utah Genome Center University of Utah Km. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0143 row: 1 column: 09
Seq primer: CACACGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 29.

FEATURES

Location/Qualifiers
1. .29

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0143109"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

```

BASE COUNT      12 a   3 c   12 g   2 t
ORIGIN

Query Match      60.0%; Score 12; DB 242; Length 29;
Best Local Similarity 75.0%; Pred. No. 7.6e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
    ||||| || |||||
Db 27 TCCTCCCTCCCTTCCTT 8

RESULT 11
LOCUS      AI355496        31 bp      mRNA      15-FEB-1999
DEFINITION qu15c07.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:1964844.3'
            similar to TR:Q60501 Q60501 P7 PREINSCRIPTION DNA. ;contains element
            TARI repetitive element ;, mRNA sequence.
ACCESSION  AI355496
VERSION     AI355496.1 GI:4095649
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 31)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 2334 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1..31
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1964844"
/clone.lib="NCI_CGAP_Ov23"
/tissue.type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/Note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.35 kb. Tumor types include: mixed
Mullerian tumor, papillary serous, clear cell, spindle
cell. All are primary tumors, metastasis positive. Life
Technologies catalog #: 11534-013"

BASE COUNT      18 a   1 c   12 g   0 t
ORIGIN

```

```

Query Match      60.0%; Score 12; DB 19; Length 31;
Best Local Similarity 75.0%; Pred. No. 7.6e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
    ||||| || |||||
Db 27 TCCTTCCTTCCTTCCTCT 8

```

```

RESULT 12
LOCUS      AZ850974        32 bp      DNA      21-FEB-2001
DEFINITION 2M0153109F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0153109 F, DNA sequence.
ACCESSION  AZ850974
VERSION     AZ850974.1 GI:13036700
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 32)
AUTHORS   Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
            Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
            M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
            and Wright, D., Weiss, R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0153 row: 1 column: 09
            Seq primer: CGTGTAAACGACGCGCCAGT
            Class: plasmid ends
            High quality sequence stop: 32.
            Location/Qualifiers
1..32
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0153109"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/Note="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

```

```

BASE COUNT      13 a   0 c   17 g   2 t

```

ORIGIN

Query Match 60.0%; Score 12; DB 250; Length 32;
 Best Local Similarity 75.0%; Pred. No. 7.7e+04;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattctt 20
 ||||| ||||| ||||| |||||
 Db 24 TCCTCCCTACTTCTTCCTT 5

RESULT 13

AI551509 34 bp mRNA EST 23-MAR-1999
 LOCUS v054h02.x1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
 DEFINITION clone IMAGE:1053747 3' similar to TR:Q33559 Q33559 NH2 TERMINUS
 UNCERTAIN ; mRNA sequence.

ACCESSION AI551509
 VERSION AI551509
 KEYWORDS EST.

SOURCE AI551509.1 GI:4483872
 house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
 B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter
 E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.

TITLE The WashU-NCI Mouse EST Project 1999
 JOURNAL Unpublished (1999)

COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:585323

This clone was previously sequenced on the 5' end only, this new
 data is from the 3' end
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 High quality sequence stop: 1.

FEATURES

source
 1. .34
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:1053747"
 /clone_lib="Barstead mouse irradiated colon MPLRB7"
 /dev_stage="8 weeks"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
 from 8 week old mouse. Colon was harvested 72 hours after
 irradiation with 1400 Gys. 1st strand cDNA was primed
 with a Not I - oligo(dT) primer
 [5'TGTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT
 T 3']; double-stranded cDNA was ligated to Eco RI
 adaptors [AATTCGATCCTTG], digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT73
 vector. Library constructed by Bob Barstead."
 4 a 9 c 1 g 20 t

BASE COUNT

ORIGIN

Query Match 60.0%; Score 12; DB 21; Length 34;
 Best Local Similarity 75.0%; Pred. No. 7.7e+04;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattctt 20
 ||||| ||||| ||||| |||||
 Db 2 TTTTCTCTTATTATTCCT 21

RESULT 14

AZ357613 34 bp DNA GSS 02-OCT-2000
 LOCUS 1M0099K18F Mouse 10kb plasmid UUCCLM library Mus musculus genomic
 DEFINITION clone UUCCLM0099K18 F, DNA sequence.

ACCESSION AZ357613
 VERSION AZ357613
 KEYWORDS GSS.

SOURCE AZ357613.1 GI:10471313
 house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS 1 (bases 1 to 34)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.,
 and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0099 row: K column: 18

Seq primer: CGTTGTAAACGCGGCCAGT

Class: plasmid ends

High quality sequence stop: 34.

FEATURES

source
 1. .34
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCCLM0099K18"
 /clone_lib="Mouse 10kb plasmid UUCCLM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adapted DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gil4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adapted vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 18 a 1 c 15 g 0 t

BASE COUNT

ORIGIN

Query Match 60.0%; Score 12; DB 242; Length 34;
 Best Local Similarity 75.0%; Pred. No. 7.7e+04;

	Matches	15;	Conservative	0;	Mismatches	5;	Indels	0;	Gaps	0;
Qy	1	ttctcccttc	ccg	ttattctt	20					
Db	22	TCCTTCCTTT	CGT	CTTTCTT	3					

Best Local Similarity 75.0%; Pred. NO. 7.8e+04;
Matches 15; Conservative 0; Mismatches 5; Indels

Qy	1	t t c t c c c t t c c g t t a t t o t t	20
Dd	32	T C C T C C C C T C T C T C T C C T	13

Qy	1	ttctcccttc	cggtatt	tctt	20
Db	32	TCCTCCCTT	CCCTCTCT	CCCTT	13

Search completed: October 2, 2001, 15:00:51
Job time: 10820 sec

RESULT	15
AZ829796/c	
LOCUS	35 bp DNA GSS 20-FEB-2001
DEFINITION	clon107H1R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0107H1 R, DNA sequence.
ACCESSION	AZ829796
VERSION	AZ829796
KEYWORDS	AZ829796.1 GI:12999704
SOURCE	GSS. house mouse. Mus musculus
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
AUTHORS	1 (bases 1 to 35) Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly, M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D., Weiss,R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	Unpublished (5)
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLCT, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0107 row: H column: 11 Seq primer: CACACAGGAACAGCTATGACC Class: plasmid ends High quality sequence stop: 35.

FEATURES	SOURCE
-----------------	---------------

	BASE COUNT	ORIGIN
1	0.0000	0.0000
2	0.0000	0.0000
3	0.0000	0.0000
4	0.0000	0.0000
5	0.0000	0.0000
6	0.0000	0.0000
7	0.0000	0.0000
8	0.0000	0.0000
9	0.0000	0.0000
10	0.0000	0.0000
11	0.0000	0.0000
12	0.0000	0.0000
13	0.0000	0.0000
14	0.0000	0.0000
15	0.0000	0.0000
16	0.0000	0.0000
17	0.0000	0.0000
18	0.0000	0.0000
19	0.0000	0.0000
20	0.0000	0.0000
21	0.0000	0.0000
22	0.0000	0.0000
23	0.0000	0.0000
24	0.0000	0.0000
25	0.0000	0.0000
26	0.0000	0.0000
27	0.0000	0.0000
28	0.0000	0.0000
29	0.0000	0.0000
30	0.0000	0.0000
31	0.0000	0.0000
32	0.0000	0.0000
33	0.0000	0.0000
34	0.0000	0.0000
35	0.0000	0.0000
36	0.0000	0.0000
37	0.0000	0.0000
38	0.0000	0.0000
39	0.0000	0.0000
40	0.0000	0.0000
41	0.0000	0.0000
42	0.0000	0.0000
43	0.0000	0.0000
44	0.0000	0.0000
45	0.0000	0.0000
46	0.0000	0.0000
47	0.0000	0.0000
48	0.0000	0.0000
49	0.0000	0.0000
50	0.0000	0.0000
51	0.0000	0.0000
52	0.0000	0.0000
53	0.0000	0.0000
54	0.0000	0.0000
55	0.0000	0.0000
56	0.0000	0.0000
57	0.0000	0.0000
58	0.0000	0.0000
59	0.0000	0.0000
60	0.0000	0.0000
61	0.0000	0.0000
62	0.0000	0.0000
63	0.0000	0.0000
64	0.0000	0.0000
65	0.0000	0.0000
66	0.0000	0.0000
67	0.0000	0.0000
68	0.0000	0.0000
69	0.0000	0.0000
70	0.0000	0.0000
71	0.0000	0.0000
72	0.0000	0.0000
73	0.0000	0.0000
74	0.0000	0.0000
75	0.0000	0.0000
76	0.0000	0.0000
77	0.0000	0.0000
78	0.0000	0.0000
79	0.0000	0.0000
80	0.0000	0.0000
81	0.0000	0.0000
82	0.0000	0.0000
83	0.0000	0.0000
84	0.0000	0.0000
85	0.0000	0.0000
86	0.0000	0.0000
87	0.0000	0.0000
88	0.0000	0.0000
89	0.0000	0.0000
90	0.0000	0.0000
91	0.0000	0.0000
92	0.0000	0.0000
93	0.0000	0.0000
94	0.0000	0.0000
95	0.0000	0.0000
96	0.0000	0.0000
97	0.0000	0.0000
98	0.0000	0.0000
99	0.0000	0.0000
100	0.0000	0.0000

Query Match . 60.0%; Score 12; DB 250; Length 35;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:46 ; Search time 417.38 Seconds
(without alignments)
9.071 Million cell updates/sec

Title: US-09-757-100B-7

Perfect score: 20
Sequence: 1 ttctcccttcggtattctt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCRUS_COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	100.0	20	3	US-09-377-310-7
2	15	75.0	15	3	US-09-377-310-27
3	13.8	69.0	41	3	US-08-813-507-49
4	12.6	63.0	30	1	US-08-480-784-75
5	12.6	63.0	30	1	US-08-483-553-75
6	12.6	63.0	30	1	US-08-487-002-75
7	12.6	63.0	30	1	US-08-483-554B-75
8	12.6	63.0	30	1	US-08-488-011B-75
9	12.6	63.0	30	4	US-08-850-727-75
10	12.6	63.0	30	5	PCT-US95-10202-75
11	12.6	63.0	30	5	PCT-US95-10203-75
12	12.6	63.0	30	5	PCT-US95-10220-75
13	12.2	61.0	22	1	US-08-701-380-4
14	12.2	61.0	22	3	US-09-032-365A-38
15	12	60.0	30	1	US-07-931-473B-35
16	12	60.0	30	1	US-07-714-131C-35
17	12	60.0	30	1	US-08-412-110-35
18	12	60.0	30	2	US-08-409-442A-35
19	12	60.0	30	2	US-08-469-609A-35
20	12	60.0	30	3	US-09-143-190-35
21	12	60.0	34	1	US-08-413-813-41
22	12	60.0	34	2	US-08-467-346-41
23	12	60.0	38	3	US-08-881-094-34
24	12	60.0	39	2	US-08-308-887A-6
25	12	60.0	39	3	US-08-881-094-6
26	12	60.0	40	1	US-08-199-507B-39
27	12	60.0	40	1	US-08-441-828-39

c	28	12	60.0	42	3	US-08-879-565-6	Sequence 6, Appli
c	29	12	60.0	50	1	US-08-133-629-6	Sequence 6, Appli
c	30	11.8	59.0	24	1	US-08-470-179-201	Sequence 201, App
c	31	11.8	59.0	27	1	US-08-758-306-426	Sequence 426, App
c	32	11.8	59.0	45	2	US-08-823-516-158	Sequence 158, App
c	33	11.6	58.0	18	2	US-08-173-489C-218	Sequence 218, App
c	34	11.6	58.0	22	1	US-08-524-757-36	Sequence 36, Appl
c	35	11.6	58.0	22	2	US-08-457-273B-18	Sequence 18, Appl
c	36	11.6	58.0	22	2	US-08-955-138-25	Sequence 25, Appl
c	37	11.6	58.0	33	3	US-08-872-417B-2	Sequence 2, Appli
c	38	11.6	58.0	35	3	US-08-872-417B-1	Sequence 1, Appli
c	39	11.6	58.0	39	3	US-09-074-357-13	Sequence 13, Appl
c	40	11.6	58.0	41	2	US-08-472-719-8	Sequence 8, Appli
c	41	11.6	58.0	41	2	US-08-472-719-9	Sequence 9, Appli
c	42	11.6	58.0	43	1	US-08-473-096-3	Sequence 3, Appli
c	43	11.6	58.0	46	1	US-08-171-389-98	Sequence 98, Appl
c	44	11.6	58.0	46	1	US-08-123-936-98	Sequence 98, Appl
c	45	11.6	58.0	46	2	US-08-475-228A-98	Sequence 98, Appl

ALIGNMENTS

RESULT 1
US-09-377-310-7
; Sequence 7, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-7

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
Db 1 ttctcccttcggtattctt 20

RESULT 2
US-09-377-310-27
; Sequence 27, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-27

Query Match 75.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ctcccttcggtatt 17

Db 1 ctcccttcggtatt 15

RESULT 3

US-08-813-507-49
; Sequence 49, Application US/08813507
; Patent No. 611416
; GENERAL INFORMATION:
; APPLICANT: Lemieux, Bertrand
; APPLICANT: Landry, Benoit S.
; APPLICANT: Sapolsky, Ronald J.
; TITLE OF INVENTION: Brassica Polymorphisms
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/813.507
FILING DATE: 07-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032.069
FILING DATE: 02-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joe
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018547-030100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 576-0200
TELEFAX: 415 576-0200
TELEX:

INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-813-507-49

Query Match 69.0%; Score 13.8; DB 3; Length 41;
Best Local Similarity 78.9%; Pred. No. 2.4e+02;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattct 19

Db 12 ttctccagtcctttattct 30

RESULT 4

US-08-480-784-75
; Sequence 75, Application US/08480784
; Patent No. 5693473

GENERAL INFORMATION:
APPLICANT: SKolnick, Mark H.
APPLICANT: Goldgar, David E.
APPLICANT: Miki, Yoshio
APPLICANT: Swenson, Jeff
APPLICANT: Ramb, Alexander
APPLICANT: Harshman, Keith D.
APPLICANT: Shattuck-Eidens, Donna M.
APPLICANT: Tavtigian, Sean V.
APPLICANT: Wiseman, Roger W.
APPLICANT: Futreal, P. Andrew
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
Susceptibility Gene
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480.784
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409.305
FILING DATE: 24-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/348.824
FILING DATE: 29-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/308.104
FILING DATE: 16-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/300.266
FILING DATE: 02-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/289.221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-08-480-784-75

Query Match 63.0%; Score 12.6; DB 1; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tctcccttcggtattctt 20

Db 8 TCTCTCTCTCTCTCTCTT 26

RESULT 5
US-08-483-553-75
; Sequence 75, Application US/08483553
; Patent No. 5709999
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio
; APPLICANT: Swenson, Jeff
; APPLICANT: Kamb, Alexander
; APPLICANT: Harshman, Keith D.
; APPLICANT: Shattuck-Eidens, Donna M.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Wiseman, Roger W.
; APPLICANT: Futreal, P. Andrew
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,553
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,305
; FILING DATE: 24-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,104
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109347
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-483-553-75

Query Match 63.08; Score 12.6; DB 1; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 tctcccttcggtattatt 20
||||| ||||| | |||||
Db 8 TCTCTCTCTCTCTCTT 26

RESULT 6
US-08-487-002-75
; Sequence 75, Application US/08487002
; Patent No. 5710001
; GENERAL INFORMATION:
; APPLICANT: Shattuck-Eidens, Donna M.
; APPLICANT: Simard, Jacques
; APPLICANT: Emi, Mitsuru
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Durocher, Francine
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,002
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,305
; FILING DATE: 24-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,104
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109347
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-487-002-75

Query Match 63.08; Score 12.6; DB 1; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 tctcccttcggtattctt 20
||||| ||||| | |||||
Db 8 TCTCTCTCTCTCTCTCTCT 26

RESULT 7
US-08-483-554B-75
; Sequence 75, Application US/08483554B
; Patent No. 5747282
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio
; APPLICANT: Swenson, Jeff
; APPLICANT: Kamb, Alexander
; APPLICANT: Harshman, Keith D.
; APPLICANT: Shattuck-Eidens, Donna M.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Wiseman, Roger W.
; APPLICANT: Futreal, P. Andrew
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483.554B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,305
; FILING DATE: 24-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,104
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109347
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens

Query Match 63.0%; Score 12.6; DB 1; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 tctcccttcggtattctt 20
||||| ||||| | |||||
Db 8 TCTCTCTCTCTCTCTCTCT 26

RESULT 8
US-08-488-011B-75
; Sequence 75, Application US/08488011B
; Patent No. 5753441
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio
; APPLICANT: Swenson, Jeff
; APPLICANT: Kamb, Alexander
; APPLICANT: Harshman, Keith D.
; APPLICANT: Shattuck-Eidens, Donna M.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Wiseman, Roger W.
; APPLICANT: Futreal, P. Andrew
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488.011B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,305
; FILING DATE: 24-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,104
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109347-09
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-488-011B-75

Query Match 63.0%; Score 12.6; DB 1; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tctcccttcctgattatttt 20
||||| ||||| |||||
Db 8 TCTCTCTTCTCTCTCTCTT 26

RESULT 9
US-08-850-727-75
; Sequence 75, Application US/08850727
; Patent No. 6162897

; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; APPLICANT: Goldgar, David E.
; APPLICANT: Miki, Yoshio
; APPLICANT: Swenson, Jeff
; APPLICANT: Kamb, Alexander
; APPLICANT: Harshman, Keith D.
; APPLICANT: Shattuck-Eidens, Donna M.
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Wiseman, Roger W.
; APPLICANT: Futreal, P. Andrew
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,727
; FILING DATE:

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,554
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,104
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109347
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-8300
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 75:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-850-727-75

Query Match 63.0%; Score 12.6; DB 4; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tctcccttcctgattatttt 20
||||| ||||| |||||
Db 8 TCTCTCTTCTCTCTCTCTT 26

RESULT 10
PCT-US95-10202-75
; Sequence 75, Application PC/TUS9510202
; GENERAL INFORMATION:
; APPLICANT: Shattuck-Eidens, Donna M.
; APPLICANT: Simard, Jacques
; APPLICANT: Emi, Mitsuru
; APPLICANT: Nakamura, Yusuke
; APPLICANT: Durocher, Francine
; TITLE OF INVENTION: In Vivo Mutations and Polymorphisms
; TITLE OF INVENTION: In the 17q-Linked Breast and Ovarian Cancer
; TITLE OF INVENTION: Susceptibility Gene
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, N.W., Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10202
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,305
; FILING DATE: 24-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08-308,104
; FILING DATE: 16-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109347
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
PCT-US95-10202-75

Query Match 63.0%; Score 12.6; DB 5; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tctcccttcggtattctt 20
||||| ||||| | |||||
Db 8 TCTCTCTCTCTCTCTT 26

RESULT 11

PCT-US95-10203-75

Sequence 75, Application PC/TUS9510203

GENERAL INFORMATION:

APPLICANT: Skolnick, Mark H.
APPLICANT: Goldgar, David E.
APPLICANT: Miki, Yoshio
APPLICANT: Swenson, Jeff
APPLICANT: Kamb, Alexander
APPLICANT: Harshman, Keith D.
APPLICANT: Shattuck-Eidens, Donna M.
APPLICANT: Tavtigian, Sean V.
APPLICANT: Wiseman, Roger W.
APPLICANT: Futreal, P. Andrew
TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
TITLE OF INVENTION: Susceptibility Gene
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:

ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/10203
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/409,305

FILING DATE: 24-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/348,824

FILING DATE: 29-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08-308,104

FILING DATE: 16-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/300,266

FILING DATE: 02-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/289,221
FILING DATE: 12-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109347
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
PCT-US95-10203-75

Query Match 63.0%; Score 12.6; DB 5; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tctcccttcggtattctt 20

||||| ||||| | |||||

Db 8 TCTCTCTCTCTCTT 26

RESULT 12

PCT-US95-10220-75

Sequence 75, Application PC/TUS9510220

GENERAL INFORMATION:

APPLICANT: Skolnick, Mark H.
APPLICANT: Goldgar, David E.
APPLICANT: Miki, Yoshio
APPLICANT: Swenson, Jeff
APPLICANT: Kamb, Alexander
APPLICANT: Harshman, Keith D.
APPLICANT: Shattuck-Eidens, Donna M.
APPLICANT: Tavtigian, Sean V.
APPLICANT: Wiseman, Roger W.
APPLICANT: Futreal, P. Andrew
TITLE OF INVENTION: Method for Diagnosing a
TITLE OF INVENTION: Predisposition for Breast and Ovarian Cancer
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:

ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/10220

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/409,305

FILING DATE: 24-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/348,824

FILING DATE: 29-NOV-1994

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08-308,104
;; FILING DATE: 16-SEP-1994
;; PRIOR APPLICATION DATA: US 08/300,266
;; FILING DATE: 02-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/289,221
;; FILING DATE: 12-AUG-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Innen, Jeffrey L.
;; REGISTRATION NUMBER: 28,957
;; REFERENCE/DOCKET NUMBER: 24884-109347
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-962-4810
;; TELEFAX: 202-962-8300
;; INFORMATION FOR SEQ ID NO: 75:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 30 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; HYPOTHETICAL: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Homo sapiens
;; PCT-US95-10220-75

Query Match 63.0%; Score 12.6; DB 5; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.5e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 tctcccttcggtattctt 20
||||| ||||| | |||||
Db 8 TCTCTCTTCCTCTCTT 26

RESULT 13
US-08-701-380-4
;; Sequence 4, Application US/08701380
;; Patent No. 5686598
;; GENERAL INFORMATION:
;; APPLICANT: NORTH, Michael
;; APPLICANT: NISHINA, Patsy
;; APPLICANT: NAGERT, Juergen
;; TITLE OF INVENTION: GENES ASSOCIATED WITH RETINAL
;; TITLE OF INVENTION: DYSTROPHIES
;; NUMBER OF SEQUENCES: 8
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
;; STREET: 4 Embarcadero Center, Suite 3400
;; CITY: San Francisco
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94111-4187
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/701,380
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SHERWOOD, Pamela J.
;; REGISTRATION NUMBER: 36,677
;; REFERENCE/DOCKET NUMBER: A-63565/PJS
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-494-8700
;; TELEFAX: 415-494-8771
;; TELEX: 910 277299

;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "primers"
;; US-08-701-380-4

Query Match 61.0%; Score 12.2; DB 1; Length 22;
Best Local Similarity 82.4%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 tctcccttcggtatttc 18
||||| ||||| | |||||
Db 2 TCTCCCTTCCTTCCTTC 18

RESULT 14
US-09-032-365A-38
;; Sequence 38, Application US/09032365A
;; Patent No. 6114502
;; GENERAL INFORMATION:
;; APPLICANT: No. 6114502th, Michael
;; APPLICANT: Nishina, Patsy
;; APPLICANT: Naggart, Juergen
;; APPLICANT: No. 6114502en-Trauth, Konrad
;; TITLE OF INVENTION: GENE FAMILY ASSOCIATED WITH
;; TITLE OF INVENTION: NEUROSENSORY DEFECTS
;; NUMBER OF SEQUENCES: 67
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Bozicevic & Reed, LLP
;; STREET: 285 Hamilton Avenue, Suite 200
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94301
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/032,365A
;; FILING DATE:
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Sherwood, Pamela J
;; REGISTRATION NUMBER: 36,677
;; REFERENCE/DOCKET NUMBER: SEQ-2CIP2
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-327-3400
;; TELEFAX: 650 327-3231
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cdna
;; US-09-032-365A-38

Query Match 61.0%; Score 12.2; DB 3; Length 22;
Best Local Similarity 82.4%; Pred. No. 1.3e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 tctcccttcggtattc 18
|||||
Db 2 TCTCCCTTCCTTCCTTC 18

RESULT 15

US-07-931-473B-35/c
; Sequence 35, Application US/07931473B
; Patent No. 5270163
; GENERAL INFORMATION:
; APPLICANT: Larry Gold
; APPLICANT: Craig Tuerk
; TITLE OF INVENTION: Nucleic Acid Ligands
; NUMBER OF SEQUENCES: 335
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beaton & Swanson, P.C.
; STREET: 4582 South Ulster Street Parkway, #403
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80237
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07931.473B
; FILING DATE: 19920817
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 850-9900
; TELEFAX: (303) 850-9401
; INFORMATION FOR SEQ ID NO: 35:
; . SEQUENCE CHARACTERISTICS:
; . LENGTH: 30 nucleotides
; . TYPE: NUCLEIC ACID
; . STRANDEDNESS: single
; . TOPOLOGY: linear
US-07-931-473B-35

Query Match 60.0%; Score 12; DB 1; Length 30;
Best Local Similarity 75.0%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 tctcccttcggtattc 20
|||||
Db 25 TTCCTTCCTTCCTTCCTTC 6

Search completed: October 2, 2001, 16:03:47
Job time: 14591 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:41 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-8

Perfect score: 20

Sequence: 1 ctatagctagctatctgtc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pil:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_om:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vil:*
- 59: gb_vil2:*
- 60: gb_vil2:*
- 61: gb_vil2:*
- 62: gb_vil2:*
- 63: gb_vil2:*
- 64: gb_vil2:*
- 65: gb_vil2:*
- 66: gb_vil2:*
- 67: gb_vil2:*
- 68: gb_vil2:*
- 69: gb_vil2:*
- 70: gb_vil2:*
- 71: gb_vil2:*
- 72: gb_vil2:*
- 73: gb_vil2:*
- 74: gb_vil2:*
- 75: gb_vil2:*
- 76: gb_vil2:*
- 77: gb_vil2:*
- 78: gb_vil2:*
- 79: gb_vil2:*
- 80: gb_vil2:*
- 81: gb_vil2:*
- 82: gb_vil2:*
- 83: gb_vil2:*
- 84: gb_vil2:*
- 85: gb_vil2:*
- 86: gb_vil2:*
- 87: gb_vil2:*
- 88: gb_vil2:*
- 89: gb_vil2:*
- 90: gb_vil2:*
- 91: gb_vil2:*
- 92: gb_vil2:*
- 93: gb_vil2:*
- 94: gb_vil2:*
- 95: gb_vil2:*
- 96: gb_vil2:*
- 97: gb_vil2:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
c 1	13.4	67.0	43	9 AR079693	Sequence AR079693
c 2	13.4	67.0	43	9 AR081223	Sequence AR081223
c 3	12.8	64.0	24	9 AG3162	Sequence AG3162
4	12.6	63.0	21	9 AR001344	Sequence AR001344
5	12.6	63.0	21	9 AR037524	Sequence AR037524
6	12.6	63.0	21	9 AR062804	Sequence AR062804
7	12.6	63.0	21	9 AR087882	Sequence AR087882
8	12.6	63.0	21	9 AR091352	Sequence AR091352

Query Match	67.0%	Score 13.4;	DB 9;	Length 43;
Best Local Similarity	93.3%	Pred. No. 4.4e+03;		
Matches 14;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps 0;				

RESULT	3
A63162	
LOCUS	A63162
	2

REFERENCE
AUTHORS
unclassified.
1 (bases 1 to 2
Buehler, T.

BASE COUNT	3 a	5
/db_xref		
/organism		

Qy 3 agatgctaggtagtctg 11
|||||

LOCUS	AR001344	Sequence	34 from
DEFINITION	AR001344		
ACCESSION	AR001344		
VERSION	AR001344	1	GT.3

SOURCE	unknown.
ORGANISM	unknown.
	unclassified.

```

VERSION AR079693.1  GI:10006434
KEYWORDS
SOURCE Unknown.

```

FEATURES
source Location/Qualifiers
1. .43
/organism="unknown"
JOURNAL Patent: US 5903/20-A 17 12 OCT 1999;

Ov 4 catagctaggatctg 18
 Best Local Similarity 93.3%; Pred. NO. 4.4e+03;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps

Db 19 GAAGCTAGGTATCTG 5

REFERENCE 1 (bases 1 to 21)
AUTHORS Kamb,A.
TITLE MTS1 beta gene
JOURNAL Patent: US 5739027-A 34 14-APR-1998;
FEATURES
LOCATION/Qualifiers
1..21
source /organism="unknown"
BASE COUNT 5 a 3 c 7 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagtgctaggtagtctgt 19
||||| ||| | |||||
Db 3 CTAGAGCGGAATTATCTGT 21

RESULT 5
AR037524 21 bp DNA PAT 29-SEP-1999
LOCUS AR037524 Sequence 34 from patent US 5801236.
DEFINITION AR037524
ACCESSION AR037524
VERSION AR037524.1 GI:5955380
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Kamb,A.
TITLE Probes for MTS1 gene and polynucleotides encoding mutant MTS1 genes
JOURNAL Patent: US 5801236-A 34 01-SEP-1998;
FEATURES
LOCATION/Qualifiers
1..21
source
BASE COUNT 5 a 3 c 7 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagtgctaggtagtctgt 19
||||| ||| | |||||
Db 3 CTAGAGCGGAATTATCTGT 21

RESULT 6
AR062804 21 bp DNA PAT 29-SEP-1999
LOCUS AR062804 Sequence 34 from patent US 5843756.
DEFINITION AR062804
ACCESSION AR062804
VERSION AR062804.1 GI:5990495
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Stone,S., Jiang,P. and Kamb,A.
TITLE Mouse MTS1 gene
JOURNAL Patent: US 5843756-A 34 01-DEC-1998;
FEATURES
LOCATION/Qualifiers
1..21
source /organism="unknown"
BASE COUNT 5 a 3 c 7 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.3e+04;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagtgctaggtagtctgt 19
||||| ||| | |||||
Db 3 CTAGAGCGGAATTATCTGT 21

RESULT 7
AR087882 21 bp DNA PAT 07-SEP-2000
LOCUS AR087882 Sequence 34 from patent US 5989815.
DEFINITION AR087882
ACCESSION AR087882
VERSION AR087882.1 GI:10014645
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 21)
AUTHORS Skolnick,M.H., Cannon-Albright,L.A. and Kamb,A.
TITLE Methods for detecting predisposition to cancer at the MTS gene
JOURNAL Patent: US 5989815-A 34 23-NOV-1999;
FEATURES
LOCATION/Qualifiers
1..21
source
BASE COUNT 5 a 3 c 7 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagtgctaggtagtctgt 19
||||| ||| | |||||
Db 3 CTAGAGCGGAATTATCTGT 21

RESULT 8
AR091352 21 bp DNA PAT 07-SEP-2000
LOCUS AR091352 Sequence 34 from patent US 5994095.
DEFINITION AR091352
ACCESSION AR091352
VERSION AR091352.1 GI:10018107
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 21)
AUTHORS Kamb,A.
TITLE MTS2 gene
JOURNAL Patent: US 5994095-A 34 30-NOV-1999;
FEATURES
LOCATION/Qualifiers
1..21
source
BASE COUNT 5 a 3 c 7 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagtgctaggtagtctgt 19
||||| ||| | |||||
Db 3 CTAGAGCGGAATTATCTGT 21

RESULT 9
I41178 21 bp DNA PAT 13-MAY-1997
LOCUS I41178 Sequence 34 from patent US 5624819.
DEFINITION I41178
ACCESSION I41178
VERSION I41178.1 GI:2081768

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 21)
AUTHORS Skolnick,M.H., Cannon-Albright,L.A. and Kamb,A.
TITLE Germ-line mutations in the MTS gene
JOURNAL Patent: US 5624819-A 34 29-APR-1997;
FEATURES Location/Qualifiers
1..21
/organism="unknown"
BASE COUNT 5 a 3 c 7 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 10; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctatagctaggtatctgt 19
||||| ||| | |||||
Db 3 CTAGAGCGCAATTACTGT 21

RESULT 10
AR012082 27 bp DNA PAT 04-DEC-1998
LOCUS AR012082
DEFINITION Sequence 6 from patent US 5763190.
ACCESSION AR012082
VERSION AR012082.1 GI:3970072
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Weiner,D.B. and Refaelli,Y.
TITLE Methods for the identification of compounds capable of inducing the nuclear translocation of a receptor complex comprising the glucocorticoid receptor type II and viral protein R interacting protein
JOURNAL Patent: US 5763190-A 6 09-JUN-1998;
FEATURES Location/Qualifiers
1..27
/organism="unknown"
BASE COUNT 6 a 6 c 8 g 7 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 27;
Best Local Similarity 78.9%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tagatgctaggtatctgtc 20
||||| ||||| |||||
Db 8 TAGTTGCTAGGCAACTGGC 26

RESULT 11
AR012083/c 27 bp DNA PAT 04-DEC-1998
LOCUS AR012083
DEFINITION Sequence 7 from patent US 5763190.
ACCESSION AR012083
VERSION AR012083.1 GI:3970073
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Weiner,D.B. and Refaelli,Y.
TITLE Methods for the identification of compounds capable of inducing the nuclear translocation of a receptor complex comprising the glucocorticoid receptor type II and viral protein R interacting protein

JOURNAL Patent: US 5763190-A 7 09-JUN-1998;
FEATURES Location/Qualifiers
1..27
/organism="unknown"
BASE COUNT 7 a 8 c 6 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 27;
Best Local Similarity 78.9%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tagatgctaggtatctgtc 20
||||| ||||| |||||
Db 24 TAGTTGCTAGGCAACTGGC 6

RESULT 12
AR030753 35 bp DNA PAT 29-SEP-1999
LOCUS AR030753
DEFINITION Sequence 2 from patent US 5861301.
ACCESSION AR030753
VERSION AR030753.1 GI:5943967
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 35)
AUTHORS Terman,B.Israel and Carrion,M.Eduardo.
TITLE Recombinant kinase insert domain containing receptor and gene encoding same
JOURNAL Patent: US 5861301-A 2 19-JAN-1999;
FEATURES Location/Qualifiers
1..35
/organism="unknown"
BASE COUNT 5 a 9 c 8 t 5 others
ORIGIN

Query Match 62.0%; Score 12.4; DB 9; Length 35;
Best Local Similarity 65.0%; Pred. No. 1.7e+04;
Matches 13; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctatagctaggtatctgtc 20
||||| ||||| |||||
Db 16 CTGCGCGCAGGCTCTGYGTC 35

RESULT 13
AX100491/c 37 bp DNA PAT 10-APR-2001
LOCUS AX100491
DEFINITION Sequence 5 from Patent WO0121783.
ACCESSION AX100491
VERSION AX100491.1 GI:13619505
KEYWORDS
SOURCE zea mays.
ORGANISM zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 37)
AUTHORS Jiao,S., Habben,J.E. and Niu,X.
TITLE Seed-preferred promoter from maize
JOURNAL Patent: WO 0121783-A 5 29-MAR-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES Location/Qualifiers
1..37
/organism="Zea mays"
/db_xref="taxon:4577"
BASE COUNT 9 a 10 c 7 g 11 t
ORIGIN

Search completed: October 2, 2001, 15:56:42
Job time: 14166 sec

Query Match 62.0%; Score 12.4; DB 10; Length 37;
Best Local Similarity 92.9%; Pred. No. 1.7e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ctagatgctaggta 14
||||| |||||
Db 37 CTAGATGATAGGTA 24

RESULT 14

AR054989 AR054989 42 bp DNA PAT 29-SEP-1999

LOCUS Sequence 2 from patent US 5837464.

DEFINITION AR054989

ACCESSION AR054989

VERSION AR054989.1 GI:5980566

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 42)

AUTHORS Capon,D. and Petropoulos,C.J.

TITLE Compositions and methods for determining anti-viral drug

susceptibility and resistance and anti-viral drug screening

JOURNAL Patent: US 5837464-A 2 17-NOV-1998;

FEATURES Location/Qualifiers

source 1..42

BASE COUNT 10 a 8 c 12 g 12 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 42;
Best Local Similarity 82.4%; Pred. No. 2.2e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ctagatgctagggtatct 17
||||| ||||| |||
Db 20 CTAGATGCTAGAGATTT 36

RESULT 15

AR036477/c AR036477 44 bp DNA PAT 29-SEP-1999

LOCUS Sequence 10 from patent US 5872218.

DEFINITION AR036477

ACCESSION AR036477

VERSION AR036477.1 GI:5953145

KEYWORDS

SOURCE

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 44)

AUTHORS Wolf,D., Tomlinson,J.E., Fretto,L.J., Giese,N.A., Escobedo,J.A. and

Williams,L.Thomas.

TITLE Human platelet-derived growth factor receptor extracellular domain

antibodies

JOURNAL Patent: US 5872218-A 10 16-FEB-1999;

FEATURES Location/Qualifiers

source 1..44

BASE COUNT 6 a 13 c 11 g 14 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 44;
Best Local Similarity 82.4%; Pred. No. 2.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gatgctagggtatctgtc 20
||||| |||||
Db 31 GAGCTAGGGGATCTGGC 15

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:00:51 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-8

Perfect score: 20

Sequence: 1 ctatagctagtagtctgtc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est75:*
137: gb_est76:*
138: gb_est77:*
139: gb_est78:*
140: gb_est79:*
141: gb_est80:*
142: gb_est81:*
143: gb_est82:*
144: gb_est83:*
145: gb_est84:*
146: gb_est85:*
147: gb_est86:*
148: gb_est87:*
149: gb_est88:*
150: gb_est89:*
151: gb_est90:*
152: gb_est99:*
153: gb_est100:*
154: gb_est101:*
155: gb_est102:*
156: gb_est103:*
157: gb_est104:*
158: gb_est105:*
159: gb_est106:*
160: gb_est67:*
161: gb_est68:*
162: gb_est69:*
163: gb_est70:*
164: gb_est71:*
165: gb_est72:*
166: gb_est73:*
167: gb_est74:*
168: gb_est91:*
169: gb_est92:*
170: gb_est93:*
171: gb_est94:*
172: gb_est95:*
173: gb_est96:*
174: gb_est97:*
175: gb_est98:*
176: em_esthum29:*
177: em_esthum30:*
178: em_esthum31:*
179: em_esthum32:*
180: em_esthum33:*
181: em_estom3:*
182: em_estpl11:*
183: em_estro21:*
184: em_estro22:*
185: em_estro23:*
186: em_htc:*
187: gb_est107:*
188: gb_est108:*
189: gb_est109:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_htc:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_other:*
207: em_gss_pln1:*
208: em_gss_pln2:*
209: em_gss_pro:*
210: em_gss_rod1:*
211: em_gss_rod2:*
212: em_gss_rod3:*
213: em_gss_rod4:*
214: em_gss_rod5:*
215: em_gss_vrt1:*
216: em_gss_vrt2:*
217: em_gss_vrt3:*
218: gb_gss1:*
219: gb_gss2:*
220: gb_gss3:*
221: gb_gss4:*
222: gb_gss5:*
223: gb_gss6:*
224: gb_gss7:*
225: gb_gss8:*
226: gb_gss9:*
227: gb_gss10:*
228: gb_gss11:*
229: gb_gss12:*
230: gb_gss13:*
231: gb_gss14:*
232: gb_gss15:*
233: gb_gss16:*
234: gb_gss17:*
235: gb_gss18:*
236: gb_gss19:*
237: gb_gss20:*
238: gb_gss21:*
239: gb_gss22:*
240: gb_gss23:*
241: gb_gss24:*
242: gb_gss25:*
243: gb_gss26:*
244: gb_gss27:*
245: gb_gss28:*
246: gb_gss29:*
247: gb_gss30:*
248: gb_gss31:*
249: gb_gss32:*
250: gb_gss33:*
251: gb_gss34:*
252: em_gss_inv4:*
253: em_gss_rod6:*
254: em_gss_rod7:*
255: em_gss_rod8:*
256: gb_gss35:*
257: gb_gss36:*
258: gb_gss37:*

```

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					TITLE	
Result No.	Score	Query Match	Length	ID	Journal COMMENT	
1	12.8	64.0	48	251	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts	
2	12.2	61.0	37	21	Unpublished (2000)	
3	12	60.0	49	12	Contact: Robert B. Weiss	
4	11.8	59.0	44	141	University of Utah Genome Center	
5	11.8	59.0	49	189	Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA	
6	11.8	59.0	50	243	Tel: 801 585 5606	
7	11.6	58.0	40	21	Fax: 801 585 7177	
8	11.6	58.0	41	244	Email: dunn@genetics.utah.edu	
9	11.6	58.0	45	250	Insert Length: 10000 Std Error: 0.00	
10	11.6	58.0	50	107	Plate: 0180 row: L column: 11	
11	11.6	58.0	50	107	Seq primer: CGTGTAAACGACGCGCAGT	
12	11.6	58.0	50	107	Class: plasmid ends	
13	11.6	58.0	50	107	High quality sequence stop: 48.	
14	11.4	57.0	36	4	Location/Qualifiers	
15	11.4	57.0	39	243	1. .48	
16	11.4	57.0	40	247	/organism="Mus musculus"	
17	11.2	56.0	33	243	/strain="C57BL/6J"	
18	11.2	56.0	33	243	/db_xref="taxon:10090"	
19	11.2	56.0	37	15	/clone="UUGC2M0180L11"	
20	11.2	56.0	46	104	/clone_lib="Mouse 10kb plasmid UUGC1M library"	
21	11.2	56.0	49	243	/sex="Male"	
22	11.2	56.0	49	246	/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"	
23	11.2	56.0	50	107	/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource	
24	11.2	56.0	50	107	(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114 gb AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."	
25	11.2	56.0	50	107	BASE COUNT 10 a 10 c 16 g 12 t	
26	11.2	56.0	50	246	ORIGIN	
27	11	55.0	26	258	Query Match 64.0%; Score 12.8; DB 251; Length 48;	
28	11	55.0	31	9	Best Local Similarity 87.5%; Pred. No. 1.7e+04;	
29	11	55.0	36	242	Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
30	11	55.0	38	249	QY 4 gatgctagggtatctgt 19	
31	11	55.0	45	244	Db 4 GGTGCTAGCTATCTGT 19	
32	11	55.0	47	246	RESULT 2	
33	11	55.0	49	2	AI545803 37 bp mRNA EST 22-MAR-1999	
34	11	55.0	50	12	LOCUS	
35	11	55.0	50	107	DEFINITION fb65e12.y1 zebrafish WashU MPIMG EST Danio rerio cDNA 5' similar to	
36	10.8	54.0	21	246	TR:Q31365 Q31365 UBA*01 CLASS I MHC MRNA. ;, mRNA sequence.	
37	10.8	54.0	30	165	AI545803	
38	10.8	54.0	31	139	ACCESSION	
39	10.8	54.0	33	143	VERSION	
40	10.8	54.0	39	242	KEYWORDS	
41	10.8	54.0	45	106	SOURCE	
42	10.8	54.0	45	106	ORGANISM	
43	10.8	54.0	46	13	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
44	10.8	54.0	49	20	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
45	10.8	54.0	50	107	1 (bases 1 to 48)	
					Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.	

Tissue Procurement: AFCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCMB33 row: k column: 09.

FEATURES

source
 1. .44
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3956864"
 /clone_lib="NIH_MGC_21"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
 Site_2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 11 a 7 c 16 g 10 t
 BASE COUNT 11 a 7 c 16 g 10 t
 ORIGIN

Query Match 59.0%; Score 11.8; DB 141; Length 44;
 Best Local Similarity 86.7%; Pred. No. 5.5e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 tagatgctaggtatc 16

Db 14 TAGGGGCTAGGTATC 28

RESULT 5

LOCUS T56680 49 bp mRNA EST 07-FEB-1995
 DEFINITION ya95b09.r1 Stratagene placenta (#937225) Homo sapiens cDNA clone
 IMAGE:69401 5' similar to similar to gb:M22995 RAS-RELATED PROTEIN
 RAP-1A (HUMAN), mRNA sequence.

ACCESSION

VERSION T56680.1 GI:658541

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 49)

Hillier L., Lennon G., Becker M., Bonaldo M.F., Chiapelli B.,

Chissoe S., Dietrich N., Dubuque T., Favello A., Gish W., Hawkins

M., Hultman M., Kucaba T., Lacy M., Le M., Le N., Mardis E., Moore

B., Morris M., Parsons J., Prange C., Rifkin L., Rohlfing T.,

Schellenberg K., Soares M.B., Tan F., Thierry-Mieg J., Trevaskis E.,

Underwood K., Wohlmann P., Waterston R., Wilson R. and Marra M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 446

High quality sequence starts: 1 High quality sequence stops: 1

Source: IMAGE Consortium, LLNL This clone is available royalty-free

through LLNL; contact the IMAGE Consortium (infoimage.llnl.gov)

for further information. Trace considered overall poor quality

Insert Length: 446 Std Error: 0.00

Seq primer: M13RPI

High quality sequence stop: 1.

FEATURES

source

1. .49

/organism="Homo sapiens"

/db_xref="GDB:490298"

/db_xref="taxon:9606"

/clone="IMAGE:69401"

/clone_lib="Stratagene placenta (#937225)"

/sex="male"

/lab_host="SOLR cells (kanamycin resistant)"

/note="Organ: placenta; Vector: pBluescript SK-; Site_1:

EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:

Oligo dT. Caucasian. Average insert size: 1.2 kb; Uni-ZAP

XR Vector; ~5' adaptor sequence: 5' GAATTCGGCAG 3' ~3'

adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'

BASE COUNT 12 a 8 c 7 g 15 t 7 others

ORIGIN

Query Match 59.0%; Score 11.8; DB 189; Length 49;

Best Local Similarity 81.2%; Pred. No. 5.6e+04;

Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 atgctaggtatctgtc 20

Db 24 ATNCTTGATCTGTC 39

RESULT 6

LOCUS AZ423820

DEFINITION 1M020322F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M020322 F, DNA sequence.

ACCESSION AZ423820

VERSION AZ423820.1 GI:10547833

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 50)

Dunn D., Aoyagi A., Barber M., Beacorn T., Duval B., Hamil C.,

Islam H., Longacre S., Mahmoud M., Meenen E., Pedersen T., Reilly

M., Rose M., Rose R., Stokes R., Tingey A., von Niederhausern A.

and Wright D., Weiss R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0203 row: p column: 22

Seq primer: CGTGTAAACGACGGCCACT

Class: plasmid ends

High quality sequence stop: 50.

Location/Qualifiers

1. .50

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M020322"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pW42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance." 9 t

BASE COUNT 15 a 15 c 11 g 9 t
 ORIGIN

Query Match 59.0%; Score 11.8; DB 243; Length 50;
 Best Local Similarity 86.7%; Pred. No. 5.6e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 agatgctaggtatctgtc 20
 ||||| |||||
 Db 49 TGTAGGTCTGTGTC 35

RESULT 7

A1499269 40 bp mRNA EST 11-MAR-1999
 LOCUS to09c08.x1 NCI-CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2178542 3'
 DEFINITION similar to TR:Q12803 Q12803 LEUKOPHYSIN.; contains element MERA0
 repetitive element ;, mRNA sequence.

ACCESSION A1499269.1 GI:4391251
 VERSION A1499269.1
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 40)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.

FEATURES

Location/Qualifiers
 1..40
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2178542"
 /clone_lib="NCI-CGAP_Ut2"
 /tissue_type="moderately-differentiated endometrial
 adenocarcinoma, 3 pooled tumors"
 /lab_hosts="DH10B"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.85 kb. Life Technologies catalog #:
 11539-012"

BASE COUNT 10 a 9 c 11 g 10 t
 ORIGIN

Query Match 58.0%; Score 11.6; DB 21; Length 40;
 Best Local Similarity 77.8%; Pred. No. 6.9e+04;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 agatgctaggtatctgtc 20
 ||||| |||||
 Db 4 AGGGCTAGGAATCTGAC 21

RESULT 8

AZ471345/c 41 bp DNA GSS 04-OCT-2000
 LOCUS 1M0285I19R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 DEFINITION Clone UUGCLM0285I19 R, DNA sequence.

ACCESSION AZ471345
 VERSION AZ471345.1 GI:10629470
 KEYWORDS GSS.
 SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 41)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL

COMMENT Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0285 row: 1 column: 19

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 41.

FEATURES

Location/Qualifiers
 1..41
 /organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGCLM0285I19"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/sex="Male"

/lab_hosts="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT

11 a 9 c 11 g 10 t

ORIGIN

Query Match 58.0%; Score 11.6; DB 244; Length 41;
 Best Local Similarity 77.8%; Pred. No. 6.9e+04;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 agatgctaggtatctgtc 20
 ||||| || |||||
 Db 41 AGATGCCAGTTTCTGCC 24

RESULT 9

AZ827171/c 45 bp DNA GSS 20-FEB-2001
 LOCUS 2M0103P19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC2M0103P19 F, DNA sequence.

ACCESSION AZ827171
 VERSION AZ827171.1 GI:12997079
 KEYWORDS GSS.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 45)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
 and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0103 row: P column: 19

Seq primer: CGTTGTAAACGACGGCCAGT

Class: plasmid ends

High quality sequence stop: 45.

Location/Qualifiers

FEATURES

source 1..45

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0103P19"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: pWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptored DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gil4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptored mouse DNA was annealed to

adaptored vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT
 ORIGIN

19 a 10 c 3 g 13 t

Query Match 58.0%; Score 11.6; DB 250; Length 45;
 Best Local Similarity 77.8%; Pred. No. 7e+04;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 tagatgctaggtatctgt 19
 ||||| ||||| |||||
 Db 26 TATATATTAGGTATATGT 9

RESULT 10

AU102832/c 50 bp mRNA EST 05-APR-2001
 LOCUS AU102832 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 DEFINITION CAS01996, mRNA sequence.

ACCESSION AU102832
 VERSION AU102832.1 GI:13552353
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 50)

AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
 ,K., Suyama,A. and Sugano,S.

TITLE Fine structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL Unpublished (2001)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano

,S. Construction and characterization of a full length-enriched and

a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source 1..50

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CAS01996"

/clone_lib="Sugano Homo sapiens cDNA library"

18 a 12 c 12 g 8 t

BASE COUNT

ORIGIN

Query Match 58.0%; Score 11.6; DB 107; Length 50;
 Best Local Similarity 77.8%; Pred. No. 7.2e+04;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 agatgctaggtatctgtc 20
 ||||| ||||| |||||
 Db 46 AGAGGTTTGGTATCTGCC 29

RESULT 11

AU102838/c 50 bp mRNA EST 05-APR-2001
 LOCUS AU102838 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 DEFINITION CAS0268, mRNA sequence.

ACCESSION AU102838
 VERSION AU102838.1 GI:13552359
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 50)

AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata

H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo, K., Suyama, A. and Sugano, S.
Fine Structural analysis of transcription start sites of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries
unpublished (2001)

COMMENT

Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshimoto-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

```

FEATURES
  source
    1..50
    Location/Qualifiers
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="CAS09268"
      /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT
  19 a 11 c 12 g 8 t
ORIGIN

```

Query Match	58.0%	Score 11.6;	DB 107;	Length 50;
Best Local Similarity	77.8%	Pred. No. 7.2e+04;		
Matches 14;	Conservative	0;	Mismatches 4;	Indels 0;
QY	3	agatgctaggtatctgtc	20	
Db	43	AGAGTTTGCTATCTGCC	26	

RESULT	12
AU102843/c	
LOCUS	50 bp mRNA EST
DEFINITION	AU102843 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone CAS11158, mRNA sequence.
ACCESSION	AU102843
VERSION	AU102843.1
KEYWORDS	GI:13552364
SOURCE	EST.
ORGANISM	human. Homo sapiens

REFERENCE
AUTHORS
 Eukaryota; Metazoa: Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo. 1 (bases 1 to 50)
 Suzuki, K., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, Y., Okubo, K., Suyama, A. and Sugano, S.
TITLE
 Fine Structural analysis of transcription start sites of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL
 Unpublished (2001)

FEATURES
source
1..50
Location/Qualifiers
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
Suzuki, Y., Yoshimoto-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
4-6-1. Shirokanedai Minato-ku, Tokyo 108-8639, Japan
Email: yszuk@ems.u-tokyo.ac.jp
Institute of Medical Science, University of Tokyo
10-50

```

Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CAS1158"
/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT      21 a      11 c      11 g      7 t
ORIGIN

Query Match      58.0%;      Score 11.6;      DB 107;      Length 50;
Best Local Similarity 77.8%;      Pred. No. 7.2e+04;
Matches 14;      Conservative 0;      Mismatches 4;      Indels 0;      Gaps 0;

```

QY	3
Db	48

RESULT	13
AU102844/C	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	

SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES SOURCE

BASE COUNT
ORIGIN

Query Match	3	48
Best Local Matches		
Qy	3	48
Db		

RESULT	14
AA275379	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	

SOURCE ORGANISM	REFERENCE AUTHORS	TITLE JOURNAL COMMENT
-----------------	-------------------	-----------------------

[illegible]

TITLE	Fine Structural analysis of transcription start sites of human MENAs using full-length enriched and 5'-end enriched cdna libraries Unpublished (2001)
JOURNAL	
COMMENT	Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp Suzuki, Y., Yoshimoto-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano , S. Construction and characterization of a full length-enriched and a 5'-end-enriched cdna library. Gene 200 (1-2), 149-156 (1997).

```

FEATURES
  source
    Location/Qualifiers
      1..50
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="cAS11168"
        /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT
  21 a 11 c 11 g 7 t

```

Query Match	58.0%;	Score 11.6;	DB 107;	Length 50;
Best Local Similarity	77.8%;	Pred. No. 7.2e+04;		
Matches 14;	Conservative	0;	Mismatches 4;	Indels 0;
QY	3	agatgctaggatctgtc	20	
Db	48	AGAGGTTTGTA	31	

RESULT	14
AA275379	
LOCUS	36 bp mRNA
DEFINITION	vcl0d01.r1 Barstead MPLRB1 Mus musculus cDNA clone IMAGE:766081 5' similar to gb:I33878 Mus musculus renal sodium phosphate (MOUSE);, mRNA sequence.
ACCESSION	AA275379
VERSION	AA275379.1 GI:1915865
KEYWORDS	EST.
SOURCE	house mouse.

SOURCE ORGANISM
mouse: mouse: mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scliuognathi; Muridae; Murinae; Mus.
1 (bases 1 to 36)

REFERENCE AUTHORS
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisell, S., Kucaba, T., Lacy, M., Le M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.

TITLE
The WashU-HMMI Mouse EST Project
Unpublished (1996)

JOURNAL COMMENT
Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:467001

Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 31.

FEATURES

source
1. 36
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:766081"
/clone_lib="Barstead MPLRBL"
/sex="mixed"
/tissue_type="Kidney"
/dev_stage="6 weeks"
/lab_host="DH108"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CATGATTCGGTACC], digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."

BASE COUNT 13 a 5 c 10 g .8 t
ORIGIN

Query Match 57.0%; Score 11.4; DB 4; Length 36;

Best Local Similarity 92.3%; Pred. No. 8.7e+04; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 tagatgctagta 14

Db 20 TACATGCTAGTA 32

RESULT 15

AZ423252

LOCUS

DEFINITION 39 bp DNA GSS 03-OCT-2000

1M0202H18F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0202H18 F, DNA sequence.

ACCESSION AZ423252

VERSION AZ423252.1 GI:10547265

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 39)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0202 row: H column: 18

Seq primer: CGTGTAAACGACGGCCACT

Class: plasmid ends

High quality sequence stop: 39.

FEATURES

source
1. 39
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0202H18"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWB42 (gil14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 9 a 8 c 3 g 19 t
ORIGIN

Query Match 57.0%; Score 11.4; DB 243; Length 39;

Best Local Similarity 92.3%; Pred. No. 8.8e+04;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 ctatgtatctgtc 20

Db 1 CTATGTATCTGTC 13

Search completed: October 2, 2001, 15:00:53

Job time: 10822 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:47 ; Search time 417.38 Seconds
(without alignments)
9.071 Million cell updates/sec

Title: US-09-757-100B-8

Perfect score: 20

Sequence: 1 ctatagctaggtagtctgtc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	20	100.0	20	3	US-09-377-310-8
2	15	75.0	15	3	US-09-377-310-28
3	15	75.0	20	3	US-09-377-310-9
C 4	13.4	67.0	43	2	US-08-850-049-17
C 5	13.4	67.0	43	2	US-08-050-478-17
C 6	13.4	67.0	43	4	US-07-858-747B-17
C 7	12.8	64.0	39	4	US-09-342-749-15
8	12.6	63.0	21	1	US-08-474-177-34
9	12.6	63.0	21	1	US-08-487-033-34
10	12.6	63.0	21	1	US-08-480-810-34
11	12.6	63.0	21	2	US-08-508-735-34
12	12.6	63.0	21	2	US-08-848-251-34
13	12.6	63.0	21	2	US-08-486-047-34
14	12.6	63.0	21	3	US-09-120-130-34
15	12.6	63.0	21	3	US-09-115-252-34
16	12.6	63.0	21	3	US-08-986-515-34
17	12.6	63.0	21	3	US-09-120-128-34
18	12.6	63.0	21	4	US-09-120-129-34
19	12.6	63.0	21	4	US-09-201-139-34
20	12.6	63.0	21	4	US-09-120-131-34
21	12.6	63.0	27	1	US-08-309-644A-6
22	12.6	63.0	27	1	US-08-309-644A-7
23	12.4	62.0	35	1	US-08-810-116-2
24	12.4	62.0	35	2	US-07-930-548A-2
25	12.2	61.0	42	2	US-08-790-963-2
26	12.2	61.0	42	4	US-09-371-774-2
C 27	12.2	61.0	44	1	US-08-168-917-10

C 28	12.2	61.0	44	2	US-08-460-510-10	Sequence 10, Appl
C 29	12.2	61.0	44	2	US-08-460-490-10	Sequence 10, Appl
C 30	12.2	61.0	44	5	PCT-US92-00730-10	Sequence 10, Appl
31	12.2	61.0	44	5	PCT-US92-10430-5	Sequence 5, Appl
32	12	60.0	15	3	US-09-377-310-29	Sequence 29, Appl
33	11.8	59.0	22	2	US-08-267-803B-21	Sequence 21, Appl
34	11.6	58.0	20	1	US-07-977-284A-74	Sequence 74, Appl
C 35	11.6	58.0	20	1	US-08-271-880A-172	Sequence 172, App
C 36	11.6	58.0	20	1	US-08-271-880A-189	Sequence 189, App
37	11.6	58.0	20	2	US-08-256-426B-74	Sequence 74, Appl
C 38	11.6	58.0	20	2	US-08-910-408-172	Sequence 172, App
C 39	11.6	58.0	20	2	US-08-910-408-189	Sequence 189, App
C 40	11.6	58.0	20	3	US-09-249-215-172	Sequence 172, App
C 41	11.6	58.0	20	3	US-09-249-215-189	Sequence 189, App
42	11.6	58.0	40	1	US-07-832-905B-69	Sequence 69, Appl
43	11.6	58.0	40	2	US-08-700-757-69	Sequence 69, Appl
C 44	11.2	56.0	22	4	US-08-876-885-21	Sequence 21, Appl
45	11.2	56.0	23	1	US-08-464-531-59	Sequence 59, Appl

ALIGNMENTS

RESULT 1
US-09-377-310-8
; Sequence 8, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-8

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctatagctaggtagtctgtc 20
Db 1 ctatagctaggtagtctgtc 20

RESULT 2
US-09-377-310-28
; Sequence 28, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

RESULT      4
US-08-850-049-17/c
; Sequence 17, Application US/08850049
; Patent No. 5965726
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT:
; APPLICANT:
; TITLE OF INVENTION: METHOD OF ELIMINATING
; TITLE OF INVENTION: INHIBITORY/INSTABILITY REGIONS OF mRNA
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850.049
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435

```

NAME: MORRY, MARY J.
REGISTRATION NUMBER: 34,398
REFERENCE/DOCKET NUMBER: 2026-4006US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
US-08-050-478-17

Query Match 67.0%; Score 13.4; DB 2; Length 43;
Best Local Similarity 93.3%; Pred. No. 1.5e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 gatgctaggtagtctg 18
||| ||||| ||||| |||||
Db 19 GAAGCTAGGTATCTG 5

RESULT 6
US-07-858-747B-17/c
Sequence 17, Application US/07858747B
Patent No. 6174666
GENERAL INFORMATION:
APPLICANT: PAVLIAKIS, GEORGE N.; FELBER, BARBARA
APPLICANT: K.
TITLE OF INVENTION: METHOD OF ELIMINATING
TITLE OF INVENTION: INHIBITORY/INSTABILITY REGIONS OF MRNA
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/858,747B
FILING DATE: 19920327
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MARY J. MORRY
REGISTRATION NUMBER: 34,398
REFERENCE/DOCKET NUMBER: 2026-4006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)758-4800
TELEFAX: (212)751-6849
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 43
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: UNKNOWN
MOLECULE TYPE: OTHER
HYPOTHETICAL: YES
FEATURE:
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: OLIGONUCLEOTIDE FOR MUTATING NT

OTHER INFORMATION: 3392-3434 OF C-FOS, HUMCFOT, ACC #V01512
US-07-858-747B-17

Query Match 67.0%; Score 13.4; DB 4; Length 43;
Best Local Similarity 93.3%; Pred. No. 1.5e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 gatgctaggtagtctg 18
||| ||||| ||||| |||||
Db 19 GAAGCTAGGTATCTG 5

RESULT 7
US-09-342-749-15/c
Sequence 15, Application US/09342749
Patent No. 6166194
GENERAL INFORMATION:
APPLICANT: Wong, Alexander K.C.
APPLICANT: Tavtigian, Sean V.
APPLICANT: Teng, David H.-F.
APPLICANT: Myriad Genetics, Inc.
TITLE OF INVENTION: TMRSS2 is a Tumor Suppressor
FILE REFERENCE: 2318-202
CURRENT APPLICATION NUMBER: US/09/342,749
CURRENT FILING DATE: 1999-06-29
EARLIER APPLICATION NUMBER: US 60/091,044
EARLIER FILING DATE: 1998-06-29
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 39
TYPE: DNA
ORGANISM: Homo sapiens
US-09-342-749-15

Query Match 64.0%; Score 12.8; DB 4; Length 39;
Best Local Similarity 87.5%; Pred. No. 3e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 atgctaggtagtctg 20
||| ||||| ||||| |||||
Db 30 ATGATAGTATCCGTC 15

RESULT 8
US-08-474-177-34
Sequence 34, Application US/08474177
Patent No. 5624819
GENERAL INFORMATION:
APPLICANT: Skolnick, Mark H.
APPLICANT: Cannon-Albright, Lisa A.
APPLICANT: Kamb, Alexander
TITLE OF INVENTION: GERMLINE MUTATIONS IN THE MTS GENE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,177
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

, APPLICATION NUMBER: US 08/251,938
, FILING DATE: 01-JUN-1994
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US 08/215,087
, FILING DATE: 18-MAR-1994
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US 08/215,086
, FILING DATE: 18-MAR-1994
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US 08/227,369
, FILING DATE: 14-APR-1994
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US 08/214,582
, FILING DATE: 18-MAR-1994
, ATTORNEY/AGENT INFORMATION:
, NAME: Ihnen, Jeffrey L.
, REGISTRATION NUMBER: 28,957
, REFERENCE/DOCKET NUMBER: 24884-109348-C
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 202-962-4810
, TELEFAX: 202-962-8300
, INFORMATION FOR SEQ ID NO: 34:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 21 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: CDNA
, HYPOTHETICAL: NO
, ANTI-SENSE: YES
, US-08-487-033-34

```

```

Query Match      63.0%; Score 12.6; DB 1; Length 21;
Best Local Similarity 78.9%; Fred. No. 3.7e-02;
Matches 15; Conservative 0; Mismatches 4; Indels

QY 1 ctagatgctagtgatctgt 19
    ||||| ||| |||||
Db 3 CTAGAGGCGAATATCTGT 21

```

```

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

APPLICATION NUMBER: US 08/215,087
FILING DATE: 18-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/215,086
FILING DATE: 18-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,369
FILING DATE: 14-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/214,582
FILING DATE: 18-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109348
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: YES
US-08-480-810-34

Query Match 63.0%; Score 12.6; DB 1; Length 21;
Best Local Similarity 78.9%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagctagggtatctgt 19
||||| || | |||||
Db 3 CTAGAGCGGAATTATCTGT 21

RESULT 11
US-08-508-735-34
Sequence 34, Application US/08508735
Patent No. 5843756
GENERAL INFORMATION:
APPLICANT: Stone, Steven
APPLICANT: Jiang, Ping
APPLICANT: Kamb, Alexander
TITLE OF INVENTION: MTS GENE AND THERAPEUTIC USE THEREOF
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/508,735
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US to be assigned
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/03316
FILING DATE: 17-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.

REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24884-109348
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4848
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: YES
US-08-508-735-34

Query Match 63.0%; Score 12.6; DB 2; Length 21;
Best Local Similarity 78.9%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagctagggtatctgt 19
||||| || | |||||
Db 3 CTAGAGCGGAATTATCTGT 21

RESULT 12
US-08-848-251-34
Sequence 34, Application US/08848251
Patent No. 5989815
GENERAL INFORMATION:
APPLICANT: Skolnick, Mark H.
APPLICANT: Cannon-Albright, Lisa A.
APPLICANT: Kamb, Alexander
TITLE OF INVENTION: GERMLINE MUTATIONS IN THE MTS GENE AND
TITLE OF INVENTION: METHOD FOR DETECTING PREDISPOSITION TO CANCER AT THE MTS
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
STREET: 1201 New York Avenue, Suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/848,251
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/474,083
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: PCT/US95/03537
FILING DATE: 17-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/251,938
FILING DATE: 01-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/215,087
FILING DATE: 18-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/215,086
FILING DATE: 18-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,369
FILING DATE: 14-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/214,582

; FILING DATE: 18-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109348-G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
US-08-848-251-34

Query Match 63.0%; Score 12.6; DB 2; Length 21;
Best Local Similarity 78.9%; Pred. NO. 3.7e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctgagtctaggtatctgt 19
||||| ||| | |||||
Db 3 CTAGAGCGCAATTATCTGT 21

RESULT 13
US-08-486-047-34
; Sequence 34, Application US/08486047
; Patent No. 5994095
; GENERAL INFORMATION:
; APPLICANT: Kamb, Alexander
; TITLE OF INVENTION: MTS2 GENE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486.047
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03316
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,938
; FILING DATE: 01-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,087
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,086
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,369
; FILING DATE: 14-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/214,582
; FILING DATE: 18-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.

; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109348-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
US-08-486-047-34

Query Match 63.0%; Score 12.6; DB 2; Length 21;
Best Local Similarity 78.9%; Pred. NO. 3.7e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctgagtctaggtatctgt 19
||||| ||| | |||||
Db 3 CTAGAGCGCAATTATCTGT 21

RESULT 14
US-09-120-130-34
; Sequence 34, Application US/09120130
; Patent No. 6037462
; GENERAL INFORMATION:
; APPLICANT: Kamb, Alexander
; TITLE OF INVENTION: MTS1 GENE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,130
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/480,810
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,938
; FILING DATE: 01-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,087
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,086
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,369
; FILING DATE: 14-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/214,582
; FILING DATE: 18-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109348
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-962-4810
; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; US-09-120-130-34

Query Match 63.0% Score 12.6; DB 3; Length 21;
Best Local Similarity 78.9%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctatagtgctaggtatctgt 19
||||| ||| |||||
Db 3 CTAGAGCGGAATTATCTGT 21

RESULT 15

US-09-115-252-34
; Sequence 34, Application US/09115252
; Patent No. 6060301
; GENERAL INFORMATION:
; APPLICANT: Kamb, Alexander
; TITLE OF INVENTION: MTS1 GENE
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Venable, Baetjer, Howard & Civiletti, LLP
; STREET: 1201 New York Avenue, Suite 1000
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/115,252
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,810
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: PCT/US95/03316
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/251,938
; FILING DATE: 01-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,087
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,086
; FILING DATE: 18-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,369
; FILING DATE: 14-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/214,582
; FILING DATE: 18-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 24884-109348
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-962-4810

; TELEFAX: 202-962-8300
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; US-09-115-252-34

Query Match 63.0% Score 12.6; DB 3; Length 21;
Best Local Similarity 78.9%; Pred. No. 3.7e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctatagtgctaggtatctgt 19
||||| ||| |||||
Db 3 CTAGAGCGGAATTATCTGT 21

Search completed: October 2, 2001, 16:03:47
Job time: 14591 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:42 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-9
Perfect score: 20
Sequence: 1 ttgtctagatgctaggtat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pl1:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_om:*
- 43: em_or:*

44: em_ov:*

45: em_pat:*

46: em_ph:*

47: em_pl:*

48: em_rod:*

49: em_sts:*

50: em_sy:*

51: em_un:*

52: em_vi:*

53: gb_sts1:*

54: gb_sts2:*

55: gb_sts3:*

56: gb_sy:*

57: gb_un:*

58: gb_vil:*

59: gb_vil2:*

60: gb_htg1:*

61: gb_htg2:*

62: gb_htg3:*

63: gb_htg4:*

64: gb_htg5:*

65: gb_htg6:*

66: gb_htg7:*

67: gb_htg8:*

68: gb_htg9:*

69: gb_htg10:*

70: gb_htg11:*

71: gb_htg12:*

72: gb_htg13:*

73: gb_htg14:*

74: gb_htg15:*

75: gb_htg16:*

76: gb_htg17:*

77: gb_htg18:*

78: gb_htg19:*

79: gb_htg20:*

80: gb_htg21:*

81: gb_htg22:*

82: gb_htg23:*

83: gb_htg24:*

84: gb_htg25:*

85: gb_pr1:*

86: gb_pr2:*

87: gb_pr3:*

88: gb_pr4:*

89: gb_pr5:*

90: gb_pr6:*

91: gb_pr7:*

92: gb_pr8:*

93: gb_pr9:*

94: gb_rol:*

95: gb_rod:*

96: gb_in4:*

97: gb_pr10:*

98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	13.6	68.0	31	10	I24252
c 2	12.4	62.0	37	10	AX100491
c 3	12.2	61.0	38	9	A25775
4	12.2	61.0	41	9	AR034285
5	12.2	61.0	41	9	AR034285 Sequence
6	12.2	61.0	41	9	AR035391
7	12.2	61.0	41	9	AR050806
8	12.2	61.0	41	9	AR053811
					AR091592 Sequence

```

9      12 60.0 26 9 A83432
10     12 60.0 33 9 AX019713
11     12 60.0 35 9 AR030753
12     12 60.0 35 9 AR075961
13     12 60.0 39 9 A42732
14     12 60.0 39 10 187238
15     12 60.0 39 10 190213
16     12 60.0 44 97 H0MA1DSC2
17     11.8 59.0 33 9 A69048
18     11.8 59.0 36 9 A71632
19     11.8 59.0 40 9 A43390
20     11.8 59.0 42 9 AR054989
21     11.6 58.0 31 9 A19662
22     11.6 58.0 37 10 AX097685
23     11.6 58.0 50 9 AR040802
24     11.4 57.0 28 9 A98431
25     11.4 57.0 33 9 AX021032
26     11.4 57.0 45 10 AX099973
27     11.2 56.0 21 10 AX097355
28     11.2 56.0 23 10 I30679
29     11.2 56.0 23 10 I46138
30     11.2 56.0 25 9 AX043430
31     11.2 56.0 25 10 E50316
32     11.2 56.0 30 9 A87635
33     11.2 56.0 31 9 AR029080
34     11.2 56.0 40 9 A17046
35     11.2 56.0 40 9 A17446
36     11.2 56.0 40 9 AR014421
37     11.2 56.0 40 10 I11864
38     11.2 56.0 41 9 A47211
39     11.2 56.0 43 10 E21659
40     11.2 56.0 48 97 S76309
41     11 55.0 20 10 E07210
42     11 55.0 21 9 A18195
43     11 55.0 21 9 A22305
44     11 55.0 22 9 AR053115
45     11 55.0 24 9 A22306

ALIGNMENTS

RESULT 1
124252 LOCUS 124252 31 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 39 from patent US 5543293.
ACCESSION I24252
VERSION I24252.1 GI:1604122
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 31)
AUTHORS Gold,L. and Tasset,D.
TITLE DNA ligands of thrombin
JOURNAL Patent: US 5543293-A 39 06-AUG-1996;
FEATURES Location/Qualifiers
source 1..31
BASE COUNT 5 a 3 c 11 g 12 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 31;
Best Local Similarity 80.0%; Pred. No. 1.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttttgctagctagctaggtat 20
||||| ||| ||||| |||
Db 2 TTTTGGTATAGCTAGGTCT 21

RESULT 2

AX100491/c AX100491 37 bp DNA PAT 10-APR-2001
LOCUS Sequence 5 from Patent WO0121783.
DEFINITION
ACCESSION AX100491
VERSION AX100491.1 GI:13619505
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
REFERENCE 1 (bases 1 to 37)
AUTHORS Jiao,S., Habben,J.E. and Niu,X.
TITLE Seed-preferred promoter from maize
JOURNAL Patent: WO 0121783-A 5 29-MAR-2001;
PIONEER HT-BRED INTERNATIONAL, INC. (US)
FEATURES Location/Qualifiers
source 1..37
/organism="Zea mays"
/db_xref="taxon:4577"
BASE COUNT 9 a 10 c 7 g 11 t
ORIGIN

Query Match 62.0%; Score 12.4; DB 10; Length 37;
Best Local Similarity 92.9%; Pred. No. 4.8e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 ctatagctagtagta 19
||||| |||||
Db 37 CTAGATGATAGGTA 24

RESULT 3
A25775/c A25775 38 bp DNA PAT 22-MAR-1995
LOCUS antisenase oligonucleotide OWB22.
ACCESSION A25775
VERSION A25775.1 GI:904750
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 38)
AUTHORS
TITLE BIOCIDAL PROTEINS
JOURNAL Patent: WO 9304586-A 8 18-MAR-1993;
FEATURES Location/Qualifiers
source 1..38
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 11 a 8 c 7 g 12 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 38;
Best Local Similarity 82.4%; Pred. No. 6.1e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 tgcctagctagctaggtat 20
||||| ||||| |||
Db 26 TGATAGAGCTAGCTAT 10

RESULT 4
AR034285 AR034285 41 bp DNA PAT 29-SEP-1999
LOCUS Sequence 8 from patent US 5869337.
ACCESSION AR034285
VERSION AR034285.1 GI:5949890
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

```

```
Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J. and
        Belshaw,P.
TITLE Regulated transcription of targeted genes and other biological
        events
JOURNAL Patent: US 5869337-A 8 09-FEB-1999;
FEATURES Location/Qualifiers
          source 1..41
BASE COUNT 7 a 13 c 9 g 12 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 41;
Best Local Similarity 82.4%; Pred. No. 6.1e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ttgtctagatgctaggt 18
   | ||||| |||||
Db 13 TCTGCTACTTGTCTAGGT 29

RESULT 5
AR035391
LOCUS AR035391 41 bp DNA 29-SEP-1999
DEFINITION Sequence 8 from patent US 5871753.
ACCESSION AR035391
VERSION AR035391.1 GI:5952059
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J.,
        Belshaw,P. and Ho,S.
TITLE Regulated transcription of targeted genes and other biological
        events
JOURNAL Patent: US 5871753-A 8 16-FEB-1999;
FEATURES Location/Qualifiers
          source 1..41
BASE COUNT 7 a 13 c 9 g 12 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 41;
Best Local Similarity 82.4%; Pred. No. 6.1e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ttgtctagatgctaggt 18
   | ||||| |||||
Db 13 TCTGCTACTTGTCTAGGT 29

RESULT 6
AR050806
LOCUS AR050806 41 bp DNA 29-SEP-1999
DEFINITION Sequence 8 from patent US 5830462.
ACCESSION AR050806
VERSION AR050806.1 GI:5974170
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J. and
        Belshaw,P.
TITLE Regulated transcription of targeted genes and other biological
        events
JOURNAL Patent: US 5830462-A 8 03-NOV-1998;
FEATURES Location/Qualifiers
          source 1..41

Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J. and
        Belshaw,P.
TITLE Regulated transcription of targeted genes and other biological
        events
JOURNAL Patent: US 5869337-A 8 09-FEB-1999;
FEATURES Location/Qualifiers
          source 1..41
BASE COUNT 7 a 13 c 9 g 12 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 41;
Best Local Similarity 82.4%; Pred. No. 6.1e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ttgtctagatgctaggt 18
   | ||||| |||||
Db 13 TCTGCTACTTGTCTAGGT 29

RESULT 7
AR053811
LOCUS AR053811 41 bp DNA 29-SEP-1999
DEFINITION Sequence 8 from patent US 5834266.
ACCESSION AR053811
VERSION AR053811.1 GI:5978673
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J. and
        Belshaw,P.
TITLE Regulated apoptosis
JOURNAL Patent: US 5834266-A 8 10-NOV-1998;
FEATURES Location/Qualifiers
          source 1..41
BASE COUNT 7 a 13 c 9 g 12 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 41;
Best Local Similarity 82.4%; Pred. No. 6.1e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ttgtctagatgctaggt 18
   | ||||| |||||
Db 13 TCTGCTACTTGTCTAGGT 29

RESULT 8
AR091592
LOCUS AR091592 41 bp DNA 07-SEP-2000
DEFINITION Sequence 8 from patent US 5994313.
ACCESSION AR091592
VERSION AR091592.1 GI:10018347
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J. and
        Belshaw,P.
TITLE Regulated apoptosis
JOURNAL Patent: US 5994313-A 8 30-NOV-1999;
FEATURES Location/Qualifiers
          source 1..41
BASE COUNT 7 a 13 c 9 g 12 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 41;
Best Local Similarity 82.4%; Pred. No. 6.1e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ttgtctagatgctaggt 18
   | ||||| |||||
Db 13 TCTGCTACTTGTCTAGGT 29

RESULT 9
AR091592
LOCUS AR091592 41 bp DNA 07-SEP-2000
DEFINITION Sequence 8 from patent US 5994313.
ACCESSION AR091592
VERSION AR091592.1 GI:10018347
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 41)
AUTHORS Crabtree,G.R., Schreiber,S.L., Spencer,D.M., Wandless,T.J. and
        Belshaw,P.
TITLE Regulated apoptosis
JOURNAL Patent: US 5994313-A 8 30-NOV-1999;
FEATURES Location/Qualifiers
          source 1..41
BASE COUNT 7 a 13 c 9 g 12 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 41;
Best Local Similarity 82.4%; Pred. No. 6.1e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ttgtctagatgctaggt 18
   | ||||| |||||
Db 13 TCTGCTACTTGTCTAGGT 29
```

```

Db 13 TCTGCTACTGCTAGCT 29

RESULT 9
A83432
LOCUS A83432 26 bp DNA PAT 21-JAN-2000
DEFINITION Sequence 18 from Patent WO9850067.
ACCESSION A83432
VERSION A83432.1 GI:6732770
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 26)
AUTHORS Goebel,W. and Demuth,A.
TITLE USE OR A SECRETION VECTOR FOR FERTILITY CONTROL BY ORAL VACCINATION
JOURNAL Patent: WO 9850067-A 18 12-NOV-1998;
GOEBEL WERNER (DE); SCHERING AG (DE)
FEATURES
source
1..26
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 6 a 5 c 7 g 8 t
ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 8e+04;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ttgtctagatgc 13
|||||
Db 4 TTTGCTAGATGC 15

RESULT 10
AX019713
LOCUS AX019713 33 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 2 from Patent WO9937807.
ACCESSION AX019713
VERSION AX019713.1 GI:10043553
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 33)
AUTHORS Hentze,M.W. and Paraskeva,E.
TITLE Method of isolation of rna-binding compounds
JOURNAL Patent: WO 9937807-A 2 29-JUL-1999;
EUROP MOLECULAR BIOLOGY LAB (DE); HENTZE MATTHIAS W (DE); PARASKEVA
EFROSYNI (DE)
FEATURES
source
1..33
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer for amplification by PCR"
BASE COUNT 7 a 4 c 4 g 18 t
ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 33;
Best Local Similarity 75.0%; Pred. NO. 7.9e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttgtctagatgctaggtat 20
|||||
Db 2 TTTTCTAGATTATTTGTAT 21

RESULT 11
AR030753
LOCUS AR030753 35 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5861301.
ACCESSION AR030753.1 GI:5943967
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 35)
AUTHORS Terman,B.Israel and Carrion,M.Eduardo.
TITLE Recombinant kinase insert domain containing receptor and gene encoding same
JOURNAL Patent: US 5861301-A 2 19-JAN-1999;
Location/Qualifiers
1..35
/organism="unknown"
BASE COUNT 5 a 9 c 8 g 8 t 5 others
ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 35;
Best Local Similarity 63.2%; Pred. No. 7.9e+04;
Matches 12; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 ttgtctagatgctaggtat 20
|||||
Db 12 KTTNCTRGCGCCAGGTCT 30

RESULT 12
AR075961
LOCUS AR075961 35 bp DNA PAT 30-AUG-2000
DEFINITION Sequence 22 from patent US 5958713.
ACCESSION AR075961
VERSION AR075961.1 GI:10002707
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 35)
AUTHORS Thastrup,O., Tullin,S.slashed.ren, Poulsen,L.kongsbak and Bj.o
slashed.rn.S.Petersen.
TITLE Method of detecting biologically active substances by using green
fluorescent protein
JOURNAL Patent: US 5958713-A 22 28-SEP-1999;
Location/Qualifiers
1..35
/organism="unknown"
BASE COUNT 16 a 7 c 3 g 9 t
ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 35;
Best Local Similarity 75.0%; Pred. No. 7.9e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttgtctagatgctaggtat 20
|||||
Db 21 TTTTGAAGAGATCTTTGGTAT 2

RESULT 13
A42732/c
LOCUS A42732 39 bp DNA PAT 06-MAR-1997
DEFINITION Sequence 64 from Patent WO9503412.
ACCESSION A42732
VERSION A42732.1 GI:2298181
KEYWORDS
SOURCE Mycobacterium fortuitum.
ORGANISM Mycobacterium fortuitum.
REFERENCE 1 (bases 1 to 39)
AUTHORS Mabilat,C. and Christen,R.

```

TITLE NUCLEOTIDE FRAGMENT OF MYCOBACTERIAL RIBOSOMAL RNA 23S, PROBES AND PRIMERS DERIVED THEREFROM, REAGENT AND METHOD FOR DETECTING SAID FRAGMENT

JOURNAL Patent: WO 9503412-A 64 02-FEB-1995;

BIO MERIEUX (FR)

COMMENT Other publication CA 2145172 950202

Other publication FR 2709310 950303.

FEATURES Location/Qualifiers

source 1. .39

/organism="Mycobacterium fortuitum"

/strain="CIP 140 410 001"

/db_xref="taxon:1766"

BASE COUNT 12 a 11 c 9 g 7 t

ORIGIN

Query Match

Best Local Similarity 60.0%; Score 12; DB 9; Length 39;

Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttttgctagatgctaggtat 20

||||| ||||| |||||

Db 25 TCTGGTAGATGTCAGGTTT 6

RESULT 14

187238/c

LOCUS 187238

DEFINITION Sequence 64 from patent US 5703217.

ACCESSION 187238

VERSION 187238.1 GI:3206956

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 39)

AUTHORS Mabilat,C. and Christen,R.

TITLE Nucleotide fragment of the 23S ribosomal RNA of mycobacteria,

derived probes and primers, reagent and detection method

JOURNAL Patent: US 5703217-A 64 30-DEC-1997;

FEATURES Location/Qualifiers

source 1. .39

/organism="unknown"

BASE COUNT 12 a 11 c 9 g 7 t

ORIGIN

Query Match

Best Local Similarity 60.0%; Score 12; DB 10; Length 39;

Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttttgctagatgctaggtat 20

||||| ||||| |||||

Db 25 TCTGGTAGATGTCAGGTTT 6

RESULT 15

190213/c

LOCUS 190213

DEFINITION Sequence 39 from patent US 5723595.

ACCESSION 190213

VERSION 190213.1 GI:3410153

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 44)

AUTHORS Thompson,G.A. and Knauf,V.C.

TITLE Plant desaturases--compositions and uses

JOURNAL Patent: US 5723595-A 39 03-MAR-1998;

FEATURES Location/Qualifiers

source 1. .44

/organism="unknown"

BASE COUNT
ORIGIN

15 a 9 c 11 g 9 t

Query Match

Best Local Similarity 60.0%; Score 12; DB 10; Length 44;

Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttttgctagatgctaggtat 20

||||| ||||| |||||

Db 42 TTTTCCTAGTTGTTAGGAT 23

Search completed: October 2, 2001, 15:56:44

Job time: 14168 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:00:53 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-9

Perfect score: 20

Sequence: 1 ttgtgtagatgctaggat 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: em_est100:*
170: em_est101:*
171: em_est102:*
172: em_est103:*
173: em_est104:*
174: em_est105:*
175: em_est106:*
176: em_est107:*
177: em_est108:*
178: em_est109:*
179: em_est110:*
180: em_est111:*
181: em_est112:*
182: em_est113:*
183: em_est114:*
184: em_est115:*
185: em_est116:*
186: em_est117:*
187: em_est118:*
188: em_est119:*
189: em_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: em_gss_fun:*
194: em_gss_hum1:*
195: em_gss_hum2:*
196: em_gss_hum3:*
197: em_gss_hum4:*
198: em_gss_hum5:*
199: em_gss_hum6:*
200: em_gss_hum7:*
201: em_gss_hum8:*
202: em_gss_hum9:*
203: em_gss_inv1:*
204: em_gss_inv2:*
205: em_gss_inv3:*
206: em_gss_inv4:*
207: em_gss_inv5:*
208: em_gss_inv6:*
209: em_gss_inv7:*
210: em_gss_inv8:*
211: em_gss_inv9:*
212: em_gss_inv10:*
213: em_gss_inv11:*
214: em_gss_inv12:*
215: em_gss_inv13:*
216: em_gss_inv14:*
217: em_gss_inv15:*
218: em_gss_inv16:*
219: em_gss_inv17:*
220: em_gss_inv18:*
221: em_gss_inv19:*
222: em_gss_inv20:*
223: em_gss_inv21:*
224: em_gss_inv22:*
225: em_gss_inv23:*
226: em_gss_inv24:*
227: em_gss_inv25:*
228: em_gss_inv26:*
229: em_gss_inv27:*
230: em_gss_inv28:*
231: em_gss_inv29:*
232: em_gss_inv30:*
233: em_gss_inv31:*
234: em_gss_inv32:*
235: em_gss_inv33:*
236: em_gss_inv34:*
237: em_gss_inv35:*
238: em_gss_inv36:*
239: em_gss_inv37:*
240: em_gss_inv38:*
241: em_gss_inv39:*
242: em_gss_inv40:*
243: em_gss_inv41:*
244: em_gss_inv42:*
245: em_gss_inv43:*
246: em_gss_inv44:*
247: em_gss_inv45:*
248: em_gss_inv46:*
249: em_gss_inv47:*
250: em_gss_inv48:*
251: em_gss_inv49:*
252: em_gss_inv50:*
253: em_gss_inv51:*
254: em_gss_inv52:*
255: em_gss_inv53:*
256: em_gss_inv54:*
257: em_gss_inv55:*
258: em_gss_inv56:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match %	Length DB	ID	Description				
C 1	12.8	64.0	23	AZ336049	AZ336049 IM0066G07				
C 2	12.6	63.0	50	AU104064	AU104064				
C 3	12.2	61.0	35	AZ841152	AZ841152 ZM0139D06				
C 4	12	60.0	49	AZ421911	AZ421911 ZM0200D22				
C 5	11.8	59.0	37	TA254H02Q	AL483435 T. brucei				
C 6	11.8	59.0	46	TA254H02Q	AA916137 og32c10.s				
C 7	11.6	58.0	37	TA383E12Q	AL497988 T. brucei				
C 8	11.6	58.0	44	BE900863	BE900863 601673838				
C 9	11.6	58.0	50	AU104334	AU104334 AU104334				
C 10	11.4	57.0	36	AA275379	AA275379 vcl0d01.r				
C 11	11.2	56.0	25	AZ396383	AZ396383 IM0160J20				
C 12	11.2	56.0	34	AZ492829	AZ492829 IM0327J11				
C 13	11.2	56.0	39	AU006827	AU006827 AU006827				
C 14	11.2	56.0	41	AZ579388	AZ579388 IM0363P15				
C 15	11.2	56.0	47	AZ588937	AZ588937 IM0397M06				
C 16	11.2	56.0	49	AA767683	AA767683 oa45b11.s				
C 17	11.2	56.0	50	AU102553	AU102553 AU102553				
C 18	11	55.0	36	AZ805982	AZ805982 ZM0067G01				
C 19	11	55.0	36	AZ827044	AZ827044 ZM0103A13				
C 20	11	55.0	37	AZ478803	AZ478803 IM0299A08				
C 21	11	55.0	40	AI146712	AI146712 qb92a01.x				
C 22	11	55.0	41	AZ358753	AZ358753 IM0101M13				
C 23	11	55.0	49	AA782046	AA782046 ai48d07.s				
C 24	11	55.0	49	AZ593107	AZ593107 IM0404O22				
C 25	11	55.0	50	AU104063	AU104063 AU104063				
C 26	10.8	54.0	21	AZ609424	AZ609424 IM0434O16				
C 27	10.8	54.0	27	AZ658164	AZ658164 IM0534I18				
C 28	10.8	54.0	30	BE276866	BE276866 G01178428				
C 29	10.8	54.0	31	BE729154	BE729154 G01561047				
C 30	10.8	54.0	33	BF026570	BF026570 G01672580				
C 31	10.8	54.0	45	AZ467641	AZ467641 IM0279I09				
C 32	10.8	54.0	47	TA388F09Q	AL498213 T. brucei				
C 33	10.6	53.0	22	AZ774731	AZ774731 ZM0004M18				
C 34	10.6	53.0	25	AZ773636	AZ773636 ZM0001O07				
C 35	10.6	53.0	30	C01989	C01989 HUMGS000401				
C 36	10.6	53.0	31	AA981706	AA981706 ua25e05.r				
C 37	10.6	53.0	33	AZ784432	AZ784432 ZM0027E05				
C 38	10.6	53.0	34	TA174FI1Q	AL474538 T. brucei				
C 39	10.6	53.0	35	C00481	C00481 HUMGS000800				
C 40	10.6	53.0	37	BF026747	BF026747 G01671956				
C 41	10.6	53.0	38	TA50D02Q	AL456084 T. brucei				
C 42	10.6	53.0	39	TA209A11P	AL479803 T. brucei				
C 43	10.6	53.0	40	AA531870	AA531870 TgESTz48				
C 44	10.6	53.0	42	AZ370282	AZ370282 IM0121C16				
C 45	10.6	53.0	46	A1948986	A1948986 wql17h03.x				

ALIGNMENTS									
RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
AZ336049/c	IM0066G07F	Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0066G07 F, DNA sequence.	AZ336049	AZ336049.1	GI:10404965	GSS.	house mouse.	1	
AZ336049	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA	AU104064	AU104064	HEP13823, mRNA sequence.	EST.	human.	1	
AU104064	50 bp	mRNA							

TITLE Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL Unpublished (2001)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
, S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

Location/Qualifiers

1..50

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HEP13823"

/clone_lib="Sugano Homo sapiens cDNA library"

8 a 15 c 15 g 12 t

BASE COUNT

ORIGIN

Query Match 63.0%; Score 12.6; DB 107; Length 50;
Best Local Similarity 78.9%; Pred. No. 2.6e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttttgcagtagctagta 19

||| ||||| |||||

Db 25 TTCTGCTAGAGGAGGTA 7

RESULT

3

A2841152/c

LOCUS

A2841152

35 bp

DNA

GSS

20-FEB-2001

DEFINITION

2M0139D06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0139D06 F, DNA sequence.

ACCESSION

A2841152

VERSION

A2841152.1

GI:13011060

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 35)

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss
University of Utah
Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0139 row: D column: 06
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 35.
Location/Qualifiers
1..35
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0139D06"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT

ORIGIN

Query Match 61.0%; Score 12.2; DB 250; Length 35;
Best Local Similarity 82.4%; Pred. No. 4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttttgcagtagctagtg 17

||| ||||| |||||

Db 20 TTTAGGTAGATGTTAGG 4

RESULT

4

A2421911/c

LOCUS

A2421911

49 bp

DNA

GSS

03-OCT-2000

DEFINITION

IM0200D22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0200D22 F, DNA sequence.

ACCESSION

A2421911

VERSION

A2421911.1

GI:10545924

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 49)

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss
University of Utah
Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0200 row: D column: 22
Seq primer: CGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 49.
Location/Qualifiers
1..49
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0200D22"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gill4732114(gb)/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 9 c 10 g 19 t
ORIGIN

Query Match 60.0%; Score 12; DB 243; Length 49;
Best Local Similarity 75.0%; Pred. No. 5.3e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttttgctagctagctagctat 20
||| | |||| |||| ||
Db 36 TTTGACAAGATCCTAGGAAT 17

RESULT 5
TA254H02Q 37 bp DNA GSS 13-DEC-2000
LOCUS
DEFINITION T. brucei sheared genomic DNA clone 254h02, reverse sequence, genomic survey sequence.
ACCESSION AL483435.1 GI:11849761
VERSION
KEYWORDS
SOURCE Trypanosoma brucei.
ORGANISM Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 37)
AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at <http://www.sanger.ac.uk/Projects/T-brucei/>.
Location/Qualifiers
1. .37
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="254h02"

BASE COUNT 8 a 8 c 10 g 11 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 258; Length 37;
Best Local Similarity 86.7%; Pred. No. 6.5e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 tttgctagctagctag 16
||| |||| |||| ||
Db 3 TTTTCTAGGTGCTAG 17

RESULT 6
AA916137 46 bp mRNA EST 14-APR-1998
LOCUS
DEFINITION OQ32c10.s1 NCI-CGAP Br7 Homo sapiens cDNA clone IMAGE:1441554.3' similar to SW:NU2M_HUMAN P03891 NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 2.; mRNA sequence.
ACCESSION AA916137
VERSION
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 46)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
unknown library type
Trace considered overall poor quality
Seq primer: -40ml3 fwd. EF from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .46
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1441554"
/clone_lib="NCI-CGAP_Br7"
/lab_host="DH10B"
/note="Organ: breast; Vector: pCMV-SPORT4; Site: 1; SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.2 kb. Life Technologies catalog #10985-018"

BASE COUNT 12 a 4 c 19 g 11 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 13; Length 46;
Best Local Similarity 86.7%; Pred. No. 6.7e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 gctagctagctagcta 19
||||| | ||||| |
Db 2 GCTAGCTCCTAGGTA 16

RESULT 7
TA383E12Q 37 bp DNA GSS 13-DEC-2000
LOCUS
DEFINITION T. brucei sheared genomic DNA clone 383e12, reverse sequence, genomic survey sequence.
ACCESSION AL497988
VERSION AL497988.1 GI:11873710
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
REFERENCE 1 (bases 1 to 37)
AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,

Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.

TITLE JOURNAL

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nhl@sanger.ac.uk

COMMENT

Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source

1..37

/organism="Trypanosoma brucei"

/strain="TREU927"

/db_xref="taxon:5691"

/clone="383e12"

11 a 5 c 6 g 15 t

BASE COUNT

ORIGIN

Query Match 58.0%; Score 11.6; DB 258; Length 37;
Best Local Similarity 77.8%; Pred. No. 8.2e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttctgtagatgtaggt 18

||||| ||| ||| ||| |||

Db 6 TTTTACTGCTGATAGGT 23

RESULT 8

LOCUS

BE900863

DEFINITION

601673838F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3956864 5',

mrna sequence.

ACCESSION

BE900863

VERSION

BE900863.1 GI:10389463

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 44)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM833 row: k column: 09.

Location/Qualifiers

1..44

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3956864"

/clone_lib="NIH_MGC_21"

/tissue_type="choriocarcinoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: placenta; Vector: pOTB7; Site:1: XhoI;

Site_2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT

ORIGIN

11 a 7 c 16 g 10 t

Query Match

Best Local Similarity

Matches 14; Conservative

0; Mismatches 4; Indels

0; Gaps 0;

Qy 3 ttgctagatgctaggtat 20

||||| ||| ||| ||| |||

Db 10 TTTATAGGGCTAGGTAT 27

RESULT 9

LOCUS

AU104334

DEFINITION

AU104334 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone

REC00112, mRNA sequence.

ACCESSION

AU104334

VERSION

AU104334.1 GI:13553855

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 50)

AUTHORS

Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata

, K., Suyama, A. and Sugano, S.

, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo

, K., Suyama, A. and Sugano, S.

Fine Structural analysis of transcription start sites of human

mRNAs using full-length enriched and 5'-end enriched cDNA libraries

Unpublished (2001)

JOURNAL

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano

, S. Construction and characterization of a full length-enriched and

a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

Location/Qualifiers

1..50

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="REC00112"

/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT

ORIGIN

5 a 8 c 20 g 17 t

Query Match

Best Local Similarity

Matches 14; Conservative

0; Mismatches 4; Indels

0; Gaps 0;

Qy 3 ttgctagatgctaggtat 20

||||| ||| ||| ||| |||

Db 10 TTGCTAGCTGTTTGCTTT 27

RESULT 10

LOCUS

AA275379

DEFINITION

AA275379 Barstead MRLB1 Mus musculus cDNA clone IMAGE:766081 5'

similar to gb:L33878 Mus musculus renal sodium phosphate (MOUSE);

mrna sequence.

ACCESSION

AA275379

VERSION

AA275379.1 GI:1915865

EST

31-MAR-1997

```

KEYWORDS EST.
SOURCE house musculus
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 36)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The WashU-HMMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:467001
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 31.
Location/Qualifiers
1. .36
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:766081"
/clone_lib="Barstead MPLRBI"
/sex="mixed"
/tissue_type="Kidney"
/dev_stage="6 weeks"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTAGCAATCTGAGTGGGCGCGCGCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(CATGATTCGGTACCT), digested with Not I and cloned into the
Not I and Eco RI sites of the modified pT7T3 vector.
Library constructed by Bob Barstead."
BASE COUNT 13 a 5 c 10 g 8 t
ORIGIN

Query Match 57.0%; Score 11.4; DB 4; Length 36;
Best Local Similarity 92.3%; Pred. No. 1e+05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7 tagatgctaggta 19
|||||
Db 20 TACATGCTAGGTA 32

RESULT 11
AZ396383/c 25 bp DNA GSS 03-OCT-2000
LOCUS IM0160J20R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
DEFINITION clone UUGCLM0160J20 R, DNA sequence.
ACCESSION AZ396383
VERSION AZ396383.1 GI:10511455
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 25)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: rdunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0160 row: J column: 20
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 25.
Location/Qualifiers
1. .25
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0160J20"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gil47321141gb1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 7-a 8 c 6 g 4 t
ORIGIN

Query Match 56.0%; Score 11.2; DB 243; Length 25;
Best Local Similarity 81.2%; Pred. No. 1.2e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 tgcctagatgctaggta 19
|||||
Db 19 TGCTGTATGCCAGGTA 4

RESULT 12
AZ492829 34 bp DNA GSS 05-OCT-2000
LOCUS IM0327J11F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
DEFINITION clone UUGCLM0327J11 F, DNA sequence.
ACCESSION AZ492829
VERSION AZ492829.1 GI:10665914
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 34)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

```



```

/db_xref="taxon:10090"
/clone="UUGC1M0363P15"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      8 a      8 c      9 g      16 t
ORIGIN

Query Match      56.0%; Score 11.2; DB 246; Length 41;
Best Local Similarity 81.2%; Pred. No. 1.3e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctgctagatgctag 16
   ||||| |||||
Db 10 TTCTGCTGATCTAG 25

RESULT 15
AZ588937/c
LOCUS      47 bp      DNA      GSS      13-DEC-2000
DEFINITION 1M0397M06R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0397M06 R, DNA sequence.
ACCESSION  AZ588937
VERSION     AZ588937.1 GI:11711127
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS   1 (bases 1 to 47)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D.,Weiss,R.
TITLE      Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL    Unpublished (2000)
COMMENT    Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0397 row: M column: 06
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 47.
Location/Qualifiers
1..47
/organism="Mus musculus"
FEATURES
source

```

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0397M06"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      16 a      8 c      11 g      12 t
ORIGIN

```

```

Query Match      56.0%; Score 11.2; DB 246; Length 47;
Best Local Similarity 81.2%; Pred. No. 1.4e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 gctagatgctaggtat 20
   ||||| |||||
Db 20 GCTAGACTAGTGT 5

```

Search completed: October 2, 2001, 15:00:54
Job time: 10823 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:44 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-10

Perfect score: 20

Sequence: 1 ggtaagcagctgccattatt 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_cm:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pli:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_cm:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vil:*
- 59: gb_vil2:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rol:*
- 95: gb_ro2:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	13.8	69.0	30	9	AR084719	AR084719 Sequence
2	13.6	68.0	30	9	A84595	A84595 Sequence 5
3	13.6	68.0	30	9	AR079892	AR079892 Sequence
C 4	13.4	67.0	18	9	A97844	A97844 Sequence 12
C 5	13.4	67.0	18	10	I44719	I44719 Sequence 45
C 6	13.4	67.0	18	10	I52125	I52125 Sequence 45
C 7	12.8	64.0	21	10	AX096725	AX096725 Sequence
C 8	12.8	64.0	26	10	E04774	E04774 Synthetic D

9 12.6 63.0 32 10 106269 106269 Sequence 27
c 10 12.4 62.0 21 10 AX094845 AX094845 Sequence
c 11 12.4 62.0 21 10 AX095462 AX095462 Sequence
c 12 12.4 62.0 39 9 AR032341 AR032341 Sequence
c 13 12.2 61.0 25 9 AR016058 AR016058 Sequence
c 14 12.2 61.0 25 9 AR025451 AR025451 Sequence
c 15 12.2 61.0 25 9 AR075530 AR075530 Sequence
c 16 12.2 61.0 25 9 AX019566 AX019566 Sequence
c 17 12.2 61.0 49 9 AR087993 AR087993 Sequence
c 18 12 60.0 25 10 E12395 E12395 PCR primer.
c 19 12 60.0 39 9 AR088301 AR088301 Sequence
c 20 12 60.0 44 10 I13034 I13034 Sequence
c 21 12 60.0 48 10 I13041 I13041 Sequence
c 22 12 60.0 48 10 I13042 I13042 Sequence
c 23 11.8 59.0 18 9 A97839 A97839 Sequence
c 24 11.8 59.0 18 9 A97845 A97845 Sequence
c 25 11.8 59.0 22 9 A97943 A97943 Sequence
c 26 11.8 59.0 39 9 AR080603 AR080603 Sequence
c 27 11.8 59.0 41 9 A81398 A81398 Sequence
c 28 11.8 59.0 41 9 A82505 A82505 Sequence
c 29 11.8 59.0 41 9 A82526 A82526 Sequence
c 30 11.8 59.0 41 9 AR064423 AR064423 Sequence
c 31 11.8 59.0 41 9 AR083774 AR083774 Sequence
c 32 11.8 59.0 41 10 I23366 I23366 Sequence
c 33 11.8 59.0 41 10 I49918 I49918 Sequence
c 34 11.8 59.0 41 10 I60381 I60381 Sequence
c 35 11.8 59.0 41 10 I85612 I85612 Sequence
c 36 11.8 59.0 49 97 S81487 S81487 T cell anti
c 37 11.8 59.0 50 9 AR032874 AR032874 Sequence
c 38 11.8 59.0 50 9 AR032875 AR032875 Sequence
c 39 11.8 59.0 50 10 I29614 I29614 Sequence
c 40 11.8 59.0 50 10 I29615 I29615 Sequence
c 41 11.8 59.0 50 10 I91288 I91288 Sequence
c 42 11.8 59.0 50 10 I91289 I91289 Sequence
c 43 11.6 58.0 22 9 AR001375 AR001375 Sequence
c 44 11.6 58.0 22 9 AR044862 AR044862 Sequence
c 45 11.6 58.0 22 9 AR052248 AR052248 Sequence

ALIGNMENTS

RESULT 1
AR084719
LOCUS AR084719 30 bp DNA
DEFINITION Sequence 6 from patent US 5981215.
ACCESSION AR084719
VERSION AR084719.1 GI:10011489
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Weissner, P.S. and Coleman, T.A.
TITLE Human cryptin growth factor
JOURNAL Patent: US 5981215-A 6 09-NOV-1999;
FEATURES Location/Qualifiers
source 1..30
BASE COUNT 11 a 7 c 3 g 9 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 9; Length 30;
Best Local Similarity 88.2%; Pred. No. 5e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 aagcagctgcattatt 20
|| ||||| |||||
Db 4 AACCAGCTGCTATTATT 20

RESULT 2

AR4595
LOCUS AR4595 30 bp DNA
DEFINITION Sequence 5 from Patent WO9845450.
ACCESSION AR4595
VERSION AR4595.1 GI:6733511
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Atkinson, E.M. and Kealey, J.T.
TITLE PURIFIED TELOMERASE
JOURNAL Patent: WO 9845450-A 5 15-OCT-1998;
GERON CORP. (US)
FEATURES Location/Qualifiers
source 1..30
/organism="unidentified"
/db_xref="taxon:32644"
modified_base 1
BASE COUNT 6 a 6 c 7 g 10 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 9; Length 30;
Best Local Similarity 80.0%; Pred. No. 6.6e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggtaagcagctgcattatt 20
|| ||||| ||||| ||
Db 6 GCCCAGCAGCTGACATTTT 25

RESULT 3
AR079892
LOCUS AR079892 30 bp DNA
DEFINITION Sequence 5 from patent US 5968506.
ACCESSION AR079892
VERSION AR079892.1 GI:10006645
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Weinrich, S.L., Atkinson, E.M. III, Lichtsteiner, S.P., Vasserot, A.P.,
Pruzan, R.A. and Kealey, J.T.
TITLE Purified telomerase
JOURNAL Patent: US 5968506-A 5 19-OCT-1999;
FEATURES Location/Qualifiers
source 1..30
BASE COUNT 6 a 6 c 7 g 10 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 9; Length 30;
Best Local Similarity 80.0%; Pred. No. 6.6e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggtaagcagctgcattatt 20
|| ||||| ||||| ||
Db 6 GCCCAGCAGCTGACATTTT 25

RESULT 4
AR7844/c
LOCUS AR7844 18 bp DNA
DEFINITION Sequence 121 from Patent WO9914377.
ACCESSION AR7844
VERSION AR7844.1 GI:6781082
KEYWORDS
SOURCE unidentified.

ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 18)
AUTHORS Quint, W. and Kleter, B.
TITLE DETECTION AND IDENTIFICATION OF HUMAN PAPILLOMAVIRUS BY PCR AND
TYPE-SPECIFIC REVERSE HYBRIDIZATION
JOURNAL Patent: WO 9914377-A 121 25-MAR-1999;
INNOGENETICS NV (BE); DELFTS DIAGNOSTIC LAB B V (NL)
FEATURES
source 1..18
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
modified_base 12
/mod_base=1
BASE COUNT 5 a 2 c 5 g 5 t 1 others
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 18;
Best Local Similarity 87.5%; Pred. No. 8.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 agcagctgccattatt 20
||||| |||||||||
Db 16 AGCANATGCCATTATT 1

RESULT 5
I44719/c I44719 18 bp DNA PAT 07-OCT-1997
LOCUS
DEFINITION Sequence 45 from patent US 5635384.
ACCESSION I44719
VERSION I44719.1 GI:2469432
KEYWORDS
SOURCE
ORGANISM
Unkown.
Unclassified.
REFERENCE 1 (bases 1 to 18)
AUTHORS Walsh, T.A., Hey, T.D. and Morgan, A.E.R.
TITLE Ribosome-inactivating proteins, inactive precursor forms thereof, a
process for making and a method of using
JOURNAL Patent: US 5635384-A 45 03-JUN-1997;
FEATURES
source 1..18
Location/Qualifiers
/organism="unknown"
BASE COUNT 5 a 4 c 4 g 5 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 10; Length 18;
Best Local Similarity 93.3%; Pred. No. 8.4e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 taagcagctgccatt 17
||||| |||||||||
Db 17 TAAGCAGCTGCAATT 3

RESULT 6
I52125/c I52125 18 bp DNA PAT 07-OCT-1997
LOCUS
DEFINITION Sequence 45 from patent US 5646026.
ACCESSION I52125
VERSION I52125.1 GI:2473326
KEYWORDS
SOURCE
ORGANISM
Unkown.
Unclassified.
REFERENCE 1 (bases 1 to 18)
AUTHORS Walsh, T.A., Hey, T.D. and Morgan, A.E.R.
TITLE Ribosome-inactivating proteins, inactive precursor forms thereof, a
process for making and a method of using
JOURNAL Patent: US 5646026-A 45 08-JUL-1997;

FEATURES
source 1..18
Location/Qualifiers
/organism="unknown"
BASE COUNT 5 a 4 c 4 g 5 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 10; Length 18;
Best Local Similarity 93.3%; Pred. No. 8.4e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 taagcagctgccatt 17
||||| |||||||||
Db 17 TAAGCAGCTGCAATT 3

RESULT 7
AX096725/c
LOCUS
DEFINITION Sequence 1903 from Patent WO0118250.
ACCESSION AX096725
VERSION AX096725.1 GI:13512979
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 21)
AUTHORS Lander, E.S., Gargill, M., Ireland, J.S., Bolk, S., Daley, G.O. and
McCarthy, J.J.
TITLE Single nucleotide polymorphisms in genes
JOURNAL Patent: WO 0118250-A 1903 15-MAR-2001;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
Pharmaceuticals, Inc. (US)
FEATURES
source 1..21
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 8 a 3 c 4 g 5 t 1 others
ORIGIN

Query Match 64.0%; Score 12.8; DB 10; Length 21;
Best Local Similarity 77.8%; Pred. No. 1.9e+04;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 gtaagcagctgccattat 19
||||| :|||||||
Db 18 GTAAAGTTRCTGCCATAAT 1

RESULT 8
E04774/c
LOCUS
DEFINITION Synthetic DNA for site directed mutagenesis of protease derived
from Bacillus NKS-21.
ACCESSION E04774
VERSION E04774.1 GI:2172970
KEYWORDS JP 1993091876-A/13.
SOURCE synthetic construct.
ORGANISM
artificial sequence.
REFERENCE 1 (bases 1 to 26)
AUTHORS Takinishi, E., Kakinuma, S., Takemoto, A., Miyoda, Y. and Fukuyama, S.
TITLE PROTEASE, ITS PRODUCTION AND USE
JOURNAL Patent: JP 1993091876-A 13 16-APR-1993;
SHOWA DENKO KK
COMMENT
OS Artificial gene
OC Artificial sequence; Genes.
OS Bacillus NKS-21
PN JP 1993091876-A/13
PD 16-APR-1993
PF 02-OCT-1991 JP 1991280313

PI TAKINISHI EIKO, KAKINUMA SHINJI, TAKEMOTO AYANO, PI MIYODA YOSHIAKI,
 PI FUKUYAMA SHIRO
 PC C12N9/54.C11D3/386.C12N1/21.C12N15/57.(C12N9/54.C12R1:07), PC
 (C12N1/21;
 PC C12R1:07);
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: NO;
 CC anti-sense: NO;

FEATURES

source
 1..26
 /organism="synthetic construct"
 /db_xref="taxon:32630"

BASE COUNT 7 a 5 c 5 g 9 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 10; Length 26;
 Best Local Similarity 87.5%; Pred. No. 1.9e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 agcagctgccattatt 20

Db 18 AGCAGCTGCAATTAAT 3

RESULT 9

LOCUS I06269 32 bp PAT 02-DEC-1994

DEFINITION Sequence 27 from Patent EP 0319052.

ACCESSION I06269

VERSION I06269.1 GI:590260

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 32)

AUTHORS Thomas Jnr,K.A. and Linemeyer,D.L.

TITLE Mutant acidic fibroblast growth factor

JOURNAL Patent: EP 0319052-A2 27 07-JUN-1989;

FEATURES Location/Qualifiers

1..32
 /organism="unknown"

BASE COUNT 5 a 8 c 8 g 11 t

ORIGIN

Query Match 63.0%; Score 12.6; DB 10; Length 32;
 Best Local Similarity 78.9%; Pred. No. 2.6e+04;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 gtaagcagctgccattatt 20

Db 7 GTTACAGCTGCAATTTT 25

RESULT 10

LOCUS AX094845/5 21 bp DNA PAT 30-MAR-2001

DEFINITION Sequence 23 from Patent WO0118250.

ACCESSION AX094845

VERSION AX094845.1 GI:13511048

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheraia; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 21)

AUTHORS Lander,E.S., Gargill,M., Ireland,J.S., Bolck,S., Daley,G.Q. and

Mccarthy,J.J.

TITLE Single nucleotide polymorphisms in genes

JOURNAL Patent: WO 0118250-A 23 15-MAR-2001;

WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium Pharmaceuticals, Inc. (US)

FEATURES
 source
 1..21
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 6 a 6 c 4 g 4 t 1 others

ORIGIN

Query Match 62.0%; Score 12.4; DB 10; Length 21;
 Best Local Similarity 81.2%; Pred. No. 3.3e+04;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 gtaagcagctgccatt 17

Db 18 GTAGGCARCTGCGATT 3

RESULT 11

LOCUS AX095462/c 21 bp DNA PAT 30-MAR-2001

DEFINITION Sequence 640 from Patent WO0118250.

ACCESSION AX095462

VERSION AX095462.1 GI:13511665

KEYWORDS human.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheraia; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 21)

AUTHORS Lander,E.S., Gargill,M., Ireland,J.S., Bolck,S., Daley,G.Q. and

Mccarthy,J.J.

TITLE Single nucleotide polymorphisms in genes

JOURNAL Patent: WO 0118250-A 640 15-MAR-2001;

WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium Pharmaceuticals, Inc. (US)

FEATURES Location/Qualifiers

1..21
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 7 a 2 c 6 g 5 t 1 others

ORIGIN

Query Match 62.0%; Score 12.4; DB 10; Length 21;
 Best Local Similarity 81.2%; Pred. No. 3.3e+04;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 agcagctgccattatt 20

Db 21 AACAGCGCGCCTTATT 6

RESULT 12

LOCUS AR032341/c 39 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 24 from patent US 5869055.

ACCESSION AR032341

VERSION AR032341.1 GI:5947946

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 39)

AUTHORS Juan,S., Lichenstein,H.S. and Wright,S.D.

TITLE Anti-inflammatory CD14 polypeptides

JOURNAL Patent: US 5869055-A 24 09-FEB-1999;

FEATURES Location/Qualifiers

1..39
 /organism="unknown"

source

BASE COUNT 7 a 13 c 13 g 6 t

ORIGIN

Query Match 62.0%; Score 12.4; DB 9; Length 39;
Best Local Similarity 92.9%; Pred. No. 3.4e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 agcagctgcccatta 18
|||||
Db 24 AGCAGCTGCCCTTA 11

RESULT 13

AR016058 LOCUS 25 bp DNA PAT 05-DEC-1998

DEFINITION Sequence 26 from patent US 5776679.

ACCESSION AR016058

VERSION AR016058.1 GI:3972335

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 25)

AUTHORS Villeponteau,B., Feng,J., Funk,W. and Andrews,W.H.

TITLE Assays for the DNA component of human telomerase

JOURNAL Patent: US 5776679-A 26 07-JUL-1998;

FEATURES Location/Qualifiers

source 1..25

BASE COUNT 8 a 8 c 6 g 3 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 25;
Best Local Similarity 82.4%; Pred. No. 4.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggtaagcagctgccatt 17
|||
Db 9 GGCCAGCAGCTGACATT 25

RESULT 14

AR025451 LOCUS 25 bp DNA PAT 05-DEC-1998

DEFINITION Sequence 22 from patent US 5798259.

ACCESSION AR025451

VERSION AR025451.1 GI:3978079

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 25)

AUTHORS Yazawa,K., Yamada,A., Kato,S. and Kondo,K.

TITLE Gene coding for eicosapentaenoic acid synthesizing enzymes and

process for production of eicosapentaenoic acid

JOURNAL Patent: US 5798259-A 22 25-AUG-1998;

FEATURES Location/Qualifiers

source 1..25

BASE COUNT 5 a 7 c 6 g 7 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 25;
Best Local Similarity 82.4%; Pred. No. 4.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 taagcagctgccattat 19
|
Db 5 TCATCAGTGCCATTAT 21

RESULT 15

AR075530 LOCUS 25 bp DNA PAT 30-AUG-2000

DEFINITION Sequence 27 from patent US 5958680.

ACCESSION AR075530

VERSION AR075530.1 GI:10002278

KEYWORDS

SOURCE

ORGANISM

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 25)

AUTHORS Villeponteau,B., Feng,J., Funk,W. and Andrews,W.H.

TITLE Mammalian telomerase

JOURNAL Patent: US 5958680-A 27 28-SEP-1999;

FEATURES Location/Qualifiers

source 1..25

BASE COUNT 8 a 8 c 6 g 3 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 25;
Best Local Similarity 82.4%; Pred. No. 4.4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggtaagcagctgccatt 17
|||
Db 9 GGCCAGCAGCTGACATT 25

Search completed: October 2, 2001, 15:56:45
Job time: 14169 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:38 ; Search time 876.95 Seconds
(without alignments)
14.320 Million cell updates/sec

Title: US-09-757-100b-10

Perfect score: 20

Sequence: 1 ggtaagcagctgccattatt 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	20	100.0	20	22	AAC65542 Human focal adhesi
c 2	16	80.0	22	20	AA336327 Sense primer used
c 3	15.2	76.0	29	16	AAT22589 Human gene signatu
4	15	75.0	15	22	AAC65562 Human focal adhesi
c 5	14.4	72.0	49	21	AAA64261 PCR primer for hum
6	14	70.0	31	20	AA06286 Human biallelic po
7	13.8	69.0	30	18	AAT51062 Human criptin grow
8	13.8	69.0	40	20	AA86412 PCR primer PDZK5.5
9	13.6	68.0	30	19	AAV63648 Antisense oligonuc
10	13.6	68.0	30	20	AAT23630 Human clone 28-1 t
c 11	13.2	66.0	21	22	AAF97138 Human gene single

c 12	13.2	66.0	29	21	AAA03899 Polymorphic fragme
c 13	13.2	66.0	30	16	AA086352 Mutagenic oligo fo
c 14	13.2	66.0	48	21	AAAL3301 PCR primer NotI.E5
c 15	13.2	66.0	48	21	AAAL3302 PCR primer ES.NotI
c 16	12.8	64.0	20	20	AA201868 PCR primer used to
c 17	12.8	64.0	21	22	AAF95258 Human gene single
c 18	12.6	64.0	21	22	AAF95875 Human gene single
c 19	12.6	63.0	29	21	AAV60352 PCR primer RP-6A u
c 20	12.6	63.0	44	18	AAV78806 Kappa light chain
c 21	12.6	63.0	44	19	AAV39247 Primer o-55i used
c 22	12.6	63.0	44	20	AAV22001 Oligonucleotide us
c 23	12.6	63.0	49	21	AAZ60353 PCR primer RP-6B u
c 24	12.4	62.0	27	20	AAZ601016 PCR primer for PGI
c 25	12.4	62.0	31	20	AA06130 Human biallelic po
c 26	12.2	61.0	24	21	AA074420 Human c-myc cDNA P
c 27	12.2	61.0	25	17	AAT34140 Primer for icosape
c 28	12.2	61.0	25	17	AAT10302 RNA component of h
c 29	12.2	61.0	25	19	AAV00513 Primer for S. putr
c 30	12.2	61.0	25	20	AAZ07280 Human telomerase R
c 31	12.2	61.0	27	19	AAV21898 Nuclease resistant
c 32	12.2	61.0	27	19	AAV21888 Nuclease resistant
c 33	12.2	61.0	30	21	AA081958 Human Meg-3 DNA pr
c 34	12.2	61.0	32	19	AAV37859 PEA toxin gene and
c 35	12.2	61.0	32	19	AAV37851 Transformed lympho
c 36	12.2	61.0	40	18	AAV57948 L-selectin truncat
c 37	12.2	61.0	45	21	AAZ88569 BYDV-PAV Illinois
c 38	12.2	61.0	49	18	AAV57947 L-selectin truncat
c 39	12.2	61.0	49	20	AA086109 Nucleic acid ligand
c 40	12	60.0	29	18	AAV96994 Human GDP-fucose p
c 41	12	60.0	29	18	AAV97584 Human GDP-fucose p
c 42	12	60.0	39	19	AAV29643 Nucleotide sequenc
c 43	12	60.0	40	19	AAV28321 24 gene cap site,
c 44	12	60.0	42	16	AA04808 B.stearothermophil
c 45	12	60.0	44	13	AA025193 TAP probe (1). Sy

ALIGNMENTS

RESULT 1
AAC65542
ID AAC65542 standard; DNA; 20 BP.
XX
AC AAC65542;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human focal adhesion kinase antisense sequence #8.
XX
KW Human; focal adhesion kinase; FAK; signal transduction; cancer;
KW embryonic development disorder; angiogenic disorder; wound healing;
KW antisense; phosphorothioate; ss.
XX
OS Homo sapiens.
XX
PN US6133031-A.
XX
PD 17-OCT-2000.
XX
PF 19-AUG-1999; 99US-0377310.
XX
PR 19-AUG-1999; 99US-0377310.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Monia BP, Gaarde WA;
XX
DR WPI; 2001-006141/01.
XX
PT New antisense compounds for inhibiting focal adhesion kinase
PT expression, especially useful for inhibiting retinal
XX neovascularization, or for diagnosing and treating e.g. colon cancer -

PS Example 2; Column 23; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
CC sequences to the human focal adhesion kinase (FAK) protein. This protein
CC is involved in integrin-mediated signal transduction, and is implicated
CC in cancer, particularly colon, breast and oral tumours, embryonic
CC development disorders, angiogenic disorders and wound healing. The
CC antisense sequences, including the one shown here, can be used in the
CC treatment of all of these.

XX Sequence 20 BP; 5 A; 4 C; 5 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggtaagcagctgccattatt 20
DB |||||

RESULT 2

AAX36327/C
ID AAX36327 standard; DNA; 22 BP.

XX AC AAX36327;

DT 21-JUL-1999 (first entry)

DE Sense primer used to amplify human FAK cDNA.

XX Human; tumorigenicity; glycosyltransferase;
KW malignancy; brain cancer cell; protein glycosylation; glioma;
KW meningioma; brain tumour; FAK; PCR primer; ss.

OS Synthetic.

XX WO9924584-A1.

PD 20-MAY-1999.

XX PF 12-NOV-1998; 98WO-US24224.

XX PR 12-NOV-1997; 97US-0969437.

XX PA (NEUR-) NEUROTHERAPEUTICS.

XX PI Moskal JR, Yamamoto H;

XX WPI; 1999-327411/27.

XX Altering tumorigenicity and malignancy of brain cancer cells

PS Example 3; Page 34; 83pp; English.

XX PCR primers AAX36327-28 were used to amplify human FAK cDNA, in the
CC course of the invention. The specification describes a method for
CC altering the tumorigenicity or malignancy of brain cancer cells by
CC changing the activity of glycosyltransferase in the cell so that
CC glycosylation of cellular proteins is modified. The method is applied
CC to glioma or meningioma, for prevention or treatment of brain tumours.
CC Measuring the level of glycosyltransferase expression in brain cells
CC is used to detect or predict their tumorigenicity.

XX Sequence 22 BP; 4 A; 7 C; 5 G; 6 T; 0 other;

Query Match 80.0%; Score 16; DB 20; Length 22;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggtaagcagctgccat 16

DB |||||

RESULT 3

AAT22589/C
ID AAT22589 standard; cDNA to mRNA; 29 BP.

XX AC AAT22589;

DT 01-OCT-1996 (first entry)

DE Human gene signature HUMGS04205.

XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.

OS Homo sapiens.

XX WO9514772-A1.

PD 01-JUN-1995.

XX PF 11-NOV-1994; 94WO-JP01916.

XX PR 12-NOV-1993; 93JP-0355504.

XX PA (MATS/) MATSUBARA K.

XX PA (OKUB/) OKUBO K.

XX PI Matsubara K, Okubo K;

XX WPI; 1995-206931/27.

XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

XX PS Claim 1; Page 1163; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared from the
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

XX Sequence 29 BP; 12 A; 6 C; 6 G; 5 T; 0 other;

Query Match 76.0%; Score 15.2; DB 16; Length 29;
Best Local Similarity 85.0%; Pred. No. 88;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggtaagcagctgccattatt 20

DB |||||

RESULT 4

AAC65562
ID AAC65562 standard; DNA; 15 BP.

XX AAC65562;
 XX AC
 XX 12-FEB-2001 (first entry)
 XX DE
 XX Human focal adhesion kinase antisense sequence #28.
 XX KW
 XX Human; focal adhesion kinase; FAK; signal transduction; cancer;
 KW embryonic development disorder; angiogenic disorder; wound healing;
 KW antisense; phosphorothioate; ss.
 XX OS
 XX Homo sapiens.
 XX PN
 XX US6133031-A.
 XX PD
 XX 17-OCT-2000.
 XX PF
 XX 19-AUG-1999; 99US-0377310.
 XX PR
 XX 19-AUG-1999; 99US-0377310.
 XX PA
 XX (ISIS-) ISIS PHARM INC.
 XX PI
 XX Monia BP, Gaarde WA;
 XX WPI; 2001-006141/01.
 XX DR
 XX New antisense compounds for inhibiting focal adhesion kinase
 PT expression, especially useful for inhibiting retinal
 PT neovascularization, or for diagnosing and treating e.g. colon cancer -
 XX Claim 15; Column 25; 30pp; English.
 XX CC
 XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.
 XX Sequence 15 BP; 4 A; 4 C; 3 G; 4 T; 0 other;
 SQ

Query Match 75.0%; Score 15; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 taagcagctgccatt 17
 Db 1 taagcagctgccatt 15
 |||||

RESULT 5
 AAA64261/c
 ID AAA64261 standard; DNA; 49 BP.
 XX AC
 XX AAA64261;
 XX 20-DEC-2000 (first entry)
 XX DE
 XX PCR primer for human fibronectin collagen-binding domain cDNA.
 XX KW
 XX Fibronectin; collagen-binding domain; sustained release; gene therapy;
 KW physiologically active polypeptide; topical retention; PCR primer;
 KW tissue regeneration; ss.
 XX OS
 XX Homo sapiens.
 XX PN
 XX WO200049159-A1.
 XX PD
 XX 24-AUG-2000.
 XX PS

PF 21-FEB-2000; 2000MO-JP00964.
 XX 19-FEB-1999; 99JP-0041913.
 PR 01-NOV-1999; 99JP-0311364.
 XX (TERU) TERUMO CORP.
 XX PA
 XX Ishikawa T, Kitajima T;
 PI WPI; 2000-565375/52.
 DR
 XX Collagen-binding active polypeptide for use in an agent for enabling
 PT topical retention or sustained release of a physiologically active
 PT peptide or physiological activity-imparting agent comprises a
 PT fibronectin peptide -
 XX Example 1; Page 124; 135pp; English.
 XX CC
 XX PCR primers AAA64261-62 were used to amplify cDNA encoding a human
 CC fibronectin collagen-binding domain. The amplified sequence was
 CC used to construct a collagen-binding physiologically active polypeptide.
 CC This polypeptide comprises a peptide from fibronectin ligated to a
 CC physiologically active peptide. The polypeptides are used in an agent
 CC for enabling topical retention or sustained release of a physiologically
 CC active peptide or physiological activity-imparting agent. They may
 CC be used in gene therapy and in tissue regeneration.
 XX Sequence 49 BP; 12 A; 16 C; 11 G; 10 T; 0 other;
 SQ

Query Match 72.0%; Score 14.4; DB 21; Length 49;
 Best Local Similarity 93.8%; Pred. No. 2.4e+02;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ggtaagcagctgccatt 16
 Db 33 GGTAACAGCTGCCAT 18
 |||||

RESULT 6
 AAX06286
 ID AAX06286 standard; DNA; 31 BP.
 XX AC
 XX AAX06286;
 XX 31-MAR-1999 (first entry)
 XX DE
 XX Human biallelic polymorphic DNA fragment SGC30775.
 XX KW
 XX Polymorphism; biallelic; paternity testing; forensic; genetic mapping;
 KW phenotypic typing; medicament; disease; marker; human; ss.
 XX OS
 XX Homo sapiens.
 XX PN
 XX WO9858529-A2.
 XX PD
 XX 30-DEC-1998.
 XX PF
 XX 22-JUN-1998; 98WO-US12930.
 XX PR
 XX 24-JUN-1997; 97US-0050594.
 XX PA
 XX (AFFY-) AFFYMETRIX INC.
 XX PI
 XX Berno A, Chee M, Fan J, Lipshutz RJ;
 XX WPI; 1999-080963/07.
 XX DR
 XX New nucleic acid segments containing polymorphic sites - used for,
 PT e.g. detecting a disease phenotype, in forensics, paternity testing
 PT or genetic mapping of phenotypic traits
 XX Claim 1; Page 18; 61pp; English.
 XX PS

XX Sequences AAX06101-X06558 represent human DNA fragments which contain
CC biallelic polymorphic markers. The base occupying the polymorphic site
CC is indicated by the appropriate IUPAC-IUB ambiguity code. These
CC fragments can be used in a method for determining polymorphic forms in
CC an individual. The invention further provides computer-readable storage
CC medium for storing data for access by an application programme being
CC executed on a data processing system. Such a method comprises a data
CC structure stored in the computer-readable storage medium, the data
CC structure including information resident in a database used by the
CC application programme and including records, each record comprising
CC information identifying a polymorphism shown in the above sequences. The
CC products and methods can be used for analysing polymorphic sites in
CC individuals for testing for the presence of a disease phenotype or in
CC forensics, paternity testing or genetic mapping of phenotypic traits.
CC They can also be used for the production of polypeptides expressed by
CC variant genes and for the production of transgenic animals. The nucleic
CC acid segments can also be used in the manufacture of medicaments for the
CC treatment or prophylaxis of diseases.
XX
SQ Sequence 31 BP; 5 A; 7 C; 6 G; 12 T; 1 other;

Query Match 70.0%; Score 14; DB 20; Length 31;
Best Local Similarity 87.5%; Pred. No. 3.7e+02;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 agcagctgccattatt 20
||||| ||||| |||||
Db 6 agcaggtgccrttatt 21

RESULT 7
AAT51062
ID AAT51062 standard; cDNA; 30 BP.

XX AAT51062;
XX
XX 05-APR-1997 (first entry)
XX Human criptin growth factor 3' primer for COS expression.

XX Criptin growth factor; CGF; angiogenesis; wound healing; vulnery;
KW muscle wastage; osteoporosis; implant fixation; tissue regeneration;
KW pancreas cancer; diagnosis; gene therapy; primer; PCR;
KW polymerase chain reaction; COS; ss.
XX
OS Synthetic.

XX WO9639420-A1.

XX 12-DEC-1996.

XX 05-JUN-1995; 95WO-US07087.

XX 05-JUN-1995; 95WO-US07087.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Coleman TA, Meissner PS;

XX WPI; 1997-043055/04.

XX New isolated human Criptin Growth Factor polypeptide - which can be
PT used to stimulate angiogenesis and develop products for use in
PT diagnosis and therapy

XX Example 3; Page 36; 52pp; English.

XX A 3' PCR primer (AAT51062) contains a PvuII site followed by 15
CC nucleotides that are the reverse complement of human criptin
CC growth factor (CGF) 3' DNA starting at the translational stop codon
CC (see also AAT51058). It was used with a 5' primer (AAT51061) to

CC amplify the CGF DNA clone deposited as ATCC 97142. The PCR product
CC was incorporated into plasmid pN346 to allow prodn. of CGF (see
CC also AAW09111) in transfected COS cells.

XX Sequence 30 BP; 11 A; 7 C; 3 G; 9 T; 0 other;

Query Match 69.0%; Score 13.8; DB 18; Length 30;
Best Local Similarity 88.2%; Pred. No. 4.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 aagcagctgccattatt 20
|| ||||| |||||
Db 4 aaccagctgctattatt 20

RESULT 8
AAX86412
ID AAX86412 standard; DNA; 40 BP.

XX AAX86412;

XX 29-SEP-1999 (first entry)

XX PCR primer PDZK5.5C used to amplify DNA encoding MMS1 protein.

XX Human; MMS1 protein; MMAC1 interacting protein; tumour suppression;
KW MMAC1 pathway; immunogen; cancer; cell neoplastic growth; PCR primer; ss.

XX Synthetic.
OS Homo sapiens.

XX WO9936566-A1.

XX 22-JUL-1999.

XX 19-JAN-1999; 99WO-US00995.

XX 20-JAN-1998; 98US-0071861.

XX (MYRI-) MYRIAD GENETICS INC.

XX Bartel PL, Tavtigian SV;

XX WPI; 1999-458472/38.

XX MMS1, an MMAC1 (tumour suppressor) interacting protein and related
PT polynucleotides

XX Example 5; Page 51; 107pp; English.

XX PCR primers AAX86368-X86423 were used to amplify DNA encoding a human
CC MMS1 protein. The PCR templates were derived from tumour cell lines,
CC and the amplicons were tested for mutations. The MMS1 protein is a
CC MMAC1 interacting protein which is involved in tumour suppression
CC activity in the MMAC1 pathway. MMS1, antigenic fragments or fusion
CC proteins of these are used as immunogens for antibody production. Primers
CC derived from MMS1 genomic clones can be used for identification of MMS1
CC genes and for synthesis by amplification of MMS1 DNA or RNA. Detecting
CC an alteration in MMS1 can be used to diagnose cancer. A germline
CC alteration in an MMS1 gene is indicative of a predisposition to cancer.
CC A somatic mutation in an MMS1 gene is indicative that the tissue is
CC cancerous. Analysis of MMAC1 and MMS1 (or PDZ domain 6 of MMS1)
CC binding interactions can be used for detection of alterations in MMAC1
CC associated with cancer. Wild-type MMS1 or a homologue can be used to
CC supply wild-type MMS1 gene function (or a substantially similar
CC function) to a cell, which has lost the gene function due to a MMS1
CC gene mutation. The gene suppresses neoplastic growth of the cell.
CC Transgenic animals having an altered MMS1 can be used as a model for
CC identifying drug candidates useful in treating cancer.

XX Sequence 40 BP; 9 A; 12 C; 7 G; 12 T; 0 other;

Query Match 69.0%; Score 13.8; DB 20; Length 40;
 Best Local Similarity 88.2%; Pred. No. 4.8e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 aagcagctgcattatt 20
 ||| ||||| |||||
 Db 16 acgcatcgcattatt 32

RESULT 9
 AAV63648
 ID AAV63648 standard; DNA; 30 BP.
 XX AC
 XX AAV63648;
 XX 15-FEB-1999 (first entry)
 XX Antisense oligonucleotide 13 for human telomerase RNA component.
 DE Human; telomerase RNA component; anticancer therapy; purification;
 KW assay; vaccine; cancer; antisense oligonucleotide; ss.
 XX Synthetic.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 modified_base 1 /*tag= a
 FT /note= "biotinylated"
 FT

XX WO9845450-A1.
 XX 15-OCT-1998.
 XX 04-APR-1997; 97WO-US06012.
 XX 04-APR-1997; 97WO-US06012.
 XX (GERO-) GERON CORP.
 XX Atkinson EM, Kealey JT, Lichtsteiner SP, Pruzan RA;
 PI Vasserot AP, Weinrich SL;
 XX WPI; 1998-594485/50.

XX Purification of telomerase on affinity material - useful for, e.g.
 diagnosis and treatment of cancer
 XX Disclosure; Page 24; 76pp; English.

XX The present sequence represents an antisense oligonucleotide
 directed against the human telomerase RNA component gene sequences.
 CC The oligonucleotide can be used as an affinity agent in the methods of
 the invention, which are used to purify human telomerase. The methods
 involve the use of several sequential steps, including the use of two
 matrices that bind molecules bearing negative charges, a matrix that
 binds molecules bearing positive charges, an affinity purification step
 and a size separation. Telomerase is a particular target of anticancer
 therapies, and is useful in assays for characterizing (pre)cancerous
 cells. Telomerase can also be used to screen for specific modulators,
 for biochemical analysis of its activity, and in preparation of
 CC antibodies. Fragments of telomerase, or nucleic acid encoding them,
 are used in vaccines, and for treating over expression of telomerase,
 particularly in cancer.

XX Sequence 30 BP; 6 A; 6 C; 8 G; 10 T; 0 other;

Query Match 68.0%; Score 13.6; DB 19; Length 30;
 Best Local Similarity 80.0%; Pred. No. 5.8e+02;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggtaagcagctgcattatt 20
 ||| ||||| ||||| ||
 Db 6 ggccagcagctgcattatt 25

RESULT 10
 AAZ23630
 ID AAZ23630 standard; DNA; 30 BP.
 XX AC
 XX AAZ23630;
 XX 07-JAN-2000 (first entry)

XX Human clone 28-1 telomerase oligonucleotide oligo-13.
 DE Telomerase; human; immune response; cancer; vaccine; treatment;
 KW disease; primer; ss.
 XX Synthetic.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 modified_base 1 /*tag= a
 FT /note= "5'-biotinylated guanosine"

XX US5968506-A.
 XX 19-OCT-1999.
 XX 04-APR-1997; 97US-0833377.
 XX 04-AUG-1995; 95US-0510736.
 XX (GERO-) GERON CORP.

XX Atkinson EM, Lichtsteiner SP, Weinrich SL, Pruzan RA, Kealey JT;
 PI Vasserot AP;
 XX WPI; 1999-590379/50.
 XX Compositions comprising human telomerase, useful for treating diseases
 associated with overexpression of telomerase e.g. cancer -
 XX Disclosure; Column 45-46; 34pp; English.

XX This invention describes a novel composition comprising human telomerase
 having at least 2000-fold (preferably at least 6000-fold) increased
 relative purity compared with crude extract of cells from
 CC adenovirus-transformed kidney cell line. The composition is useful for
 CC eliciting an immune response in animals and may therefore be used as a
 CC vaccine for treating diseases associated with the overexpression of
 CC telomerase e.g. cancer. AAZ23628-223637 represent oligonucleotides used
 CC in the isolation of human clone 28-1 which contains a fragment of the
 CC human telomerase described in the method of the invention.

XX Sequence 30 BP; 6 A; 6 C; 8 G; 10 T; 0 other;

Query Match 68.0%; Score 13.6; DB 20; Length 30;
 Best Local Similarity 80.0%; Pred. No. 5.8e+02;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggtaagcagctgcattatt 20
 ||| ||||| ||||| ||
 Db 6 ggccagcagctgcattatt 25

RESULT 11
 AAF97138/c
 ID AAF97138 standard; DNA; 21 BP.
 XX AC
 XX AAF97138;

XX
DT 06-JUN-2001 (first entry)
XX
DE Human gene single nucleotide polymorphism #1899.
XX
KW Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW polymorphism; vascular disease; coronary artery disease; forensics;
KW myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW pulmonary embolism; paternity test; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Variation replace(11,T)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO200118250-A2.
XX
PD 15-MAR-2001.
XX
PF 07-SEP-2000; 2000WO-US24503.
XX
PR 10-SEP-1999; 99US-0153357.
PR 26-JUL-2000; 2000US-0220947.
PR 16-AUG-2000; 2000US-0225724.
XX
PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lander ES, Gargill M, Ireland JS, Boik S, Daley GO, McCarthy JJ;
XX
DR WPI; 2001-226749/23.
XX
PT Nucleic acids comprising single nucleotide polymorphisms useful in
PT applications such as forensics, paternity testing, medicine, genetic
PT analysis and phenotype correlations to diseases such as diabetes and
PT atherosclerosis -
XX
PS Examples; Page 177; 242pp; English.
XX
CC The present invention provides a method of diagnosing a vascular disease
CC in an individual, involving determining the sequence at various
CC polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC genes. The sequences at a number of polymorphic sites are also provided
CC in the specification. In particular, the method can be used in the
CC diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC disease, stroke, peripheral vascular diseases, venous thromboembolism
CC and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC useful in forensics, paternity testing, genetic analysis and phenotype
CC correlations to diseases. The present sequence is an example of one of
CC the human gene SNPs shown in the specification.
XX
SQ Sequence 21 BP; 8 A; 4 C; 4 G; 5 T; 0 other;

Query Match 66.0%; Score 13.2; DB 22; Length 21;
Best Local Similarity 83.3%; Pred. No. 8.9e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 gtaagcagctgcattat 19
||||| ||||| |||
Db 18 GTAAGTCTGCTGCTAAT 1

RESULT 12
AAA03899
ID AAA03899 standard; DNA; 29 BP.
XX
AC AAA03899;
XX
DT 22-MAY-2000 (first entry)
XX

DE Polymorphic fragment of hypertension associated gene AGT.
XX
KW Polymorphism; hypertension; agammaglobulinemia; diabetes insipidus;
KW Leach-Nyhan syndrome; muscular dystrophy; Wiskott-Aldrich syndrome;
KW Fabry's disease; familial hypercholesterolemia; hereditary spherocytosis;
KW polycystic kidney disease; von Willebrands disease; hereditary telangiectasia;
KW tuberous sclerosis; hereditary hemorrhagica telangiectasia;
KW familial colonic polyposis; osteogenesis imperfecta; porphyria;
KW Ehlers-Danlos syndrome; ss.
XX
OS Homo sapiens.
XX
PN EP955382-A2.
XX
PD 10-NOV-1999.
XX
PF 07-MAY-1999; 99EP-0250150.
XX
PR 07-MAY-1999; 98US-0084641.
PR 03-MAY-1999; 99US-0304232.
XX
PA (AFY-) AFFYMETRIX INC.
PA (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI Fan JB, Chakravarti A, Haluska MK;
XX
DR WPI; 2000-107928/10.
XX
PT Novel nucleic acids containing polymorphisms used in the diagnosis of
PT hypertension -
XX
PS Claim 1; Page 19; 53pp; English.
XX
CC The invention provides polymorphic fragments of genes associated with
CC hypertension. The nucleic acids including the polymorphic sites can be
CC used as probes or primers for expressing variant proteins. Detection of
CC the polymorphisms is useful in designing prophylactic and therapeutic
CC regimes customized to underlying abnormalities. The polymorphisms can be
CC used for association studies for hypertension, and in hypertension
CC diagnostic assays. Where the polymorphisms have strong correlation with
CC hypertension, within a gene, they are likely to have a causative role in
CC hypertension. This information can be used to find the precise role of a
CC polymorphism in the disease, and this can be used to identify potential
CC drugs which combat the disease. The polymorphisms can be tested for
CC association with other diseases e.g. agammaglobulinemia, diabetes
CC insipidus, Leach-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich
CC syndrome, Fabry's disease, familial hypercholesterolemia, polycystic
CC kidney disease, hereditary spherocytosis, von Willebrands disease,
CC tuberous sclerosis, hereditary hemorrhagica telangiectasia, familial
CC colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, and
CC acute intermittent porphyria. The polymorphic forms can also be used in
CC forensics to identify individuals.
XX
SQ Sequence 29 BP; 5 A; 4 C; 12 G; 7 T; 1 other;

Query Match 66.0%; Score 13.2; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 9.3e+02;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggtaagcagctgcattatt 20
||||| ||||| |||
Db 2 gggaagaagctgcygtgtt 21

RESULT 13
AAQ86352/C
ID AAQ86352 standard; DNA; 30 BP.
XX
AC AAQ86352;
XX
DT 29-DEC-1995 (first entry)
XX

DE Mutagenic oligo for human protein S variant I425A/I426A.
 XX Protein S; PS; vitamin K-dependent protein; mutagenic oligo; ss.
 KW Synthetic.
 OS
 XX Key Location/Qualifiers
 FH misc_difference 13..14 /tag= a
 FT /note= "nts for introducing desired AA substn."
 FT misc_difference 16..17 /tag= b
 FT /note= "nts for introducing desired AA substn."
 FT
 XX US5405946-A.
 XX 11-APR-1995.
 XX 02-DEC-1992; 92US-0985691.
 XX 02-DEC-1992; 92US-0985691.
 XX (SCRI) SCRIPPS RES INST.
 XX Bertina R, Bouma BN, Griffin JH;
 XX WPI; 1995-154630/20.
 XX
 XX New recombinant protein S variants - having reduced C4b binding
 PT protein binding activity and anticoagulant activity for treating
 PT thrombosis
 XX
 PS Example; Column 18; 24pp; English.
 XX
 CC For preparing a protein S (PS) expression vector, partial cDNAs
 CC coding for human protein S were first isolated as described by Ploos
 CC van Amstel et al., FEBS Lett., 222:186-190 (1987) from a pUC9 human
 CC liver cDNA library. The cDNA sequence is given in Q86348. The PS nt
 CC sequence is also listed in GenBank having the accession number Y00692.
 CC The mRNA encodes a preprotein having 676 AAs. After post-
 CC translation processing the corresp. translated mature PS consists of
 CC 635 AAs as given in R72350. The AA sequence is also listed in
 CC GenBank having the accession no. A26157. PS can be modified without
 CC significant loss of anticoagulant activity by introducing one or
 CC more mutations in the region between residues 425 and 432 to reduce
 CC significantly or eliminate the ability of PS to bind C4BP. A
 CC variant protein S is claimed having AA residue substitutions
 CC K429E, I425A/I426A and K432E (AA residue nos. corresp. to R72350).
 CC Oligos Q86349-Q86350 were used to produce mutations
 CC in the first disulfide loop in the sex hormone binding globulin-like
 CC domain of PS.
 XX
 SQ Sequence 30 BP; 16 A; 4 C; 7 G; 3 T; 0 Other;

Query Match 66.0%; Score 13.2; DB 16; Length 30;
 Best Local Similarity 83.3%; Pred. No. 9.3e+02;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 taagcagctgcccattatt 20
 | | | | | | | | | | | | | |
 Db 20 TGAGCAGCTTCCTTTATT 3

RESULT 14
 AAAL3301
 ID AAAL3301 standard; cDNA; 48 BP.
 AC
 AC AAAL3301;
 XX
 XX 24-JUL-2000 (first entry)
 DT
 DE PCR primer NotI.E5 used in the production of vectors for trans-splicing.

XX

XX Recombinant nucleic acid production; combinatorial gene library;
 KW ordered gene assembly; trans-splicing; tissue plasminogen activator;
 KW vector production; PCR primer; ss.
 XX Synthetic.
 OS
 XX WO200017342-A2.
 XX 30-MAR-2000.
 XX 21-SEP-1999; 99WO-US21929.
 XX 21-SEP-1998; 98US-0101328.
 XX 20-SEP-1999; 99US-0101328.
 XX (UYBO-) UNIV BOSTON.
 XX Jarrell KA, Mulcheeva S, Donahue W;
 XX WPI; 2000-303208/26.
 XX In vivo production of nucleic acid, useful e.g. for producing
 PT combinatorial gene libraries or ribozymes, by trans-splicing two RNAs
 PT containing exon and intron component
 XX
 PS Example 4; Page 49; 186pp; English.
 XX
 CC This sequence represents a PCR primer used in the construction of
 CC trans-splicing vectors to illustrate the invention. The present invention
 CC relates to the in vivo production of recombinant nucleic acid sequences.
 CC The method comprises expressing in a cell, two transcripts, one
 CC containing a first exon and first intron component, and a second
 CC transcript comprising a second intron component and a second exon.
 CC Transcript 1 and transcript 2 are allowed to trans-splice, forming a
 CC product containing exon 1 and exon 2, but not the intron components. The
 CC invention makes use of the ability of intronic sequences derived from
 CC group I or group II introns to mediate specific cleavage and ligation of
 CC discontinuous nucleic acid molecules. The method is used to produce
 CC recombinant nucleic acids, their products, or ribozymes in vivo. The
 CC method is preferably used for the preparation of combinatorial gene
 CC libraries in which the order and composition of exons are random. The
 CC method is also used for ordered gene assembly, or for the assembly of
 CC genes ordered at some exons but randomized at others. New genes can be
 CC selected rapidly and efficiently, and a very wide range of exons may be
 CC trapped.
 XX
 SQ Sequence 48 BP; 16 A; 12 C; 10 G; 10 T; 0 other;

Query Match 66.0%; Score 13.2; DB 21; Length 48;
 Best Local Similarity 83.3%; Pred. No. 1e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gtaagcagctgcccattat 19
 | | | | | | | | | | | | | |
 Db 17 gaaagcgccgcccattat 34

RESULT 15
 AAAL3302/c
 ID AAAL3302 standard; cDNA; 48 BP.
 XX
 AC AAAL3302;
 XX

DT 24-JUL-2000 (first entry)
 XX
 DE PCR primer E5.NotI used in the production of vectors for trans-splicing.
 XX Recombinant nucleic acid production; combinatorial gene library;
 KW ordered gene assembly; trans-splicing; tissue plasminogen activator;
 KW vector production; PCR primer; ss.
 XX

OS Synthetic.
XX
PN WO200017342-A2.
XX
PD 30-MAR-2000.
XX
PF 21-SEP-1999; 99WO-US21929.
XX
PR 21-SEP-1998; 98US-0101328.
PR 20-SEP-1999; 99US-0101328.
XX
PA (UYBO-) UNIV BOSTON.
XX
PI Jarrell KA, Mulcheeva S, Donahue W;
XX
DR WPI; 2000-303208/26.
XX
PT In vivo production of nucleic acid, useful e.g. for producing
PT combinatorial gene libraries or ribozymes, by trans-splicing two RNAs
PT containing exon and intron component
XX
PS Example 4; Page 49; 186pp; English.
XX
CC This sequence represents a PCR primer used in the construction of
CC trans-splicing vectors to illustrate the invention. The present invention
CC relates to the in vivo production of recombinant nucleic acid sequences.
CC The method comprises expressing in a cell, two transcripts, one
CC containing a first exon and first intron component, and a second
CC transcript comprising a second intron component and a second exon.
CC Transcript 1 and transcript 2 are allowed to trans-splice, forming a
CC product containing exon 1 and exon 2, but not the intron components. The
CC invention makes use of the ability of intronic sequences derived from
CC group I or group II introns to mediate specific cleavage and ligation of
CC discontinuous nucleic acid molecules. The method is used to produce
CC recombinant nucleic acids, their products, or ribozymes in vivo. The
CC method is preferably used for the preparation of combinatorial gene
CC libraries in which the order and composition of exons are random. The
CC method is also used for ordered gene assembly, or for the assembly of
CC genes ordered at some exons but randomized at others. New genes can be
CC selected rapidly and efficiently, and a very wide range of exons may be
CC trapped.
XX
SQ Sequence 48 BP; 10 A; 10 C; 12 G; 16 T; 0 other;

Query Match 66.0%; Score 13.2; DB 21; Length 48;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 gtaagcagctgccattat 19
| | | | | | | | | |
Db 32 GAAAGCGCGCGCCATTAT 15

Search completed: October 2, 2001, 16:18:39
Job time: 15483 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:00:54 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-10

Perfect score: 20

Sequence: 1 ggtaagcagctgccattatt 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est25:*
- 26: gb_est26:*
- 27: gb_est27:*
- 28: gb_est28:*
- 29: gb_est29:*
- 30: gb_est30:*
- 31: gb_est31:*
- 32: gb_est32:*
- 33: gb_est33:*
- 34: gb_est34:*
- 35: gb_est35:*
- 36: gb_est36:*
- 37: gb_est37:*
- 38: gb_est38:*
- 39: gb_est39:*
- 40: gb_est40:*
- 41: gb_est41:*
- 42: gb_est42:*
- 43: gb_est43:*
- 44: gb_est44:*
- 45: gb_est45:*
- 46: gb_est46:*
- 47: gb_est47:*
- 48: gb_est48:*
- 49: gb_est49:*
- 50: gb_est50:*
- 51: gb_est51:*
- 52: gb_est52:*
- 53: gb_est53:*
- 54: gb_est54:*
- 55: gb_est55:*
- 56: gb_est56:*
- 57: gb_est57:*
- 58: gb_est58:*
- 59: gb_est59:*
- 60: gb_est60:*
- 61: gb_est61:*
- 62: gb_est62:*
- 63: gb_est63:*
- 64: gb_est64:*
- 65: gb_est65:*
- 66: gb_est66:*
- 67: gb_est67:*
- 68: gb_est68:*
- 69: gb_est69:*
- 70: gb_est70:*
- 71: gb_est71:*
- 72: gb_est72:*
- 73: gb_est73:*
- 74: gb_est74:*
- 75: gb_est75:*
- 76: gb_est76:*
- 77: gb_est77:*
- 78: gb_est78:*
- 79: gb_est79:*
- 80: gb_est80:*
- 81: gb_est81:*
- 82: gb_est82:*
- 83: gb_est83:*
- 84: gb_est84:*
- 85: gb_est85:*
- 86: gb_est86:*
- 87: gb_est87:*
- 88: gb_est88:*
- 89: gb_est89:*
- 90: gb_est90:*
- 91: gb_est91:*
- 92: gb_est92:*
- 93: gb_est93:*
- 94: gb_est94:*
- 95: gb_est95:*
- 96: gb_est96:*
- 97: gb_est97:*
- 98: gb_est98:*
- 99: gb_est99:*
- 100: gb_est100:*
- 101: gb_est101:*
- 102: gb_est102:*
- 103: gb_est103:*
- 104: gb_est104:*
- 105: gb_est105:*
- 106: gb_est106:*
- 107: gb_est107:*
- 108: gb_est108:*
- 109: gb_est109:*
- 110: gb_est110:*
- 111: gb_est111:*
- 112: gb_est112:*
- 113: gb_est113:*
- 114: gb_est114:*
- 115: gb_est115:*
- 116: gb_est116:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est33:*
- 111: gb_est34:*
- 112: gb_est35:*
- 113: gb_est36:*
- 114: gb_est37:*
- 115: gb_est38:*
- 116: gb_est39:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES										JOURNAL COMMENT									
Result No.	Score	Query Match %	Length	DB ID	Description	mrnas using full-length enriched and 5'-end enriched cDNA libraries													
C 1	15.2	76.0	50	107	AU103462	Unpublished (2001)													
C 2	13.2	66.0	41	247	AZ634745	Contact: Yutaka Suzuki													
C 3	12.4	62.0	45	12	AA794942	Department of Virology													
C 4	12.2	61.0	28	258	TA355F03Q	Institute of Medical Science, University of Tokyo													
C 5	12.2	61.0	30	258	TA272C09Q	4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan													
C 6	12.2	61.0	40	12	AA789284	Email: yusuzuki@ims.u-tokyo.ac.jp													
C 7	11.8	59.0	31	249	AZ798762	Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano													
C 8	11.8	59.0	33	241	AZ3111284	S. Construction and characterization of a full length-enriched and													
C 9	11.8	59.0	41	12	AA829551	a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).													
C 10	11.8	59.0	45	188	TI7561	Location/Qualifiers													
C 11	11.8	59.0	50	10	AA662608	1.50													
C 12	11.8	59.0	50	107	AU102666	/organism="Homo sapiens"													
C 13	11.8	59.0	50	107	AU103485	/db_xref="taxon:9606"													
C 14	11.8	59.0	50	107	AU103486	/clone="HEP14884"													
C 15	11.8	59.0	50	107	AU103487	/clone_lib="Sugano Homo sapiens cDNA library"													
C 16	11.8	59.0	50	107	AU103488	9 a 15 c 9 g 17 t													
C 17	11.8	59.0	50	107	AU103489	BASE COUNT													
C 18	11.8	59.0	50	107	AU103493	ORIGIN													
C 19	11.6	58.0	40	106	AU006965	Query Match 76.0%; Score 15.2; DB 107; Length 50;													
C 20	11.6	58.0	43	244	AZ476191	Best Local Similarity 85.0%; Pred. No. 1.7e+03;													
C 21	11.6	58.0	44	249	AZ797074	Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;													
C 22	11.6	58.0	48	244	AZ492766	QY 1 ggtaagcagctgccattatt 20													
C 23	11.6	58.0	49	13	AA914273														
C 24	11.6	58.0	50	107	AU105748	Db 41 GGTGAGCAGCTGCTATAATT 22													
C 25	11.4	57.0	31	156	D18234	RESULT 2													
C 26	11.4	57.0	34	249	AZ785014	AZ634745/c													
C 27	11.4	57.0	47	143	BF036368	LOCUS													
C 28	11.4	57.0	50	107	AU103473	DEFINITION													
C 29	11.2	56.0	19	244	AZ481008	1M0490012R Mouse 10kb plasmid UUGC1M library Mus musculus genomic													
C 30	11.2	56.0	20	249	AZ768476	clonase UUGC1M0490012 R, DNA sequence.													
C 31	11.2	56.0	22	246	AZ598225	ACCESSION													
C 32	11.2	56.0	33	147	BF346987	AZ634745													
C 33	11.2	56.0	39	250	AZ834606	VERSION													
C 34	11.2	56.0	43	241	AZ329485	GSS.													
C 35	11.2	56.0	50	20	A1476044	SOURCE													
C 36	11.2	56.0	50	107	AU107071	house mouse.													
C 37	11.2	56.0	50	107	AU107430	Mus musculus													
C 38	11.2	56.0	50	107	AU107431	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;													
C 39	11.2	56.0	50	173	BG077090	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.													
C 40	11	55.0	26	156	D21048	1 (bases 1 to 41)													
C 41	11	55.0	36	243	AZ425719	Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,													
C 42	11	55.0	36	244	AZ482058	Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly													
C 43	11	55.0	40	247	AZ644893	M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.													
C 44	11	55.0	41	258	TA386D04P	and Wright, D., Weiss, R.													
C 45	11	55.0	46	192	AK019062	Mouse whole genome scaffolding with paired end reads from 10kb													
ALIGNMENTS										plasmid inserts									
RESULT 1										Unpublished (2000)									
AU103462/c										Contact: Robert B. Weiss									
LOCUS										University of Utah									
DEFINITION										Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT									
ACCESSION										84112, USA									
VERSION										Tel: 801 585 5606									
KEYWORDS										Fax: 801 585 7177									
SOURCE										Email: ddunn@genetics.utah.edu									
ORGANISM										Insert Length: 10000 Std Error: 0.00									
REFERENCE										Plate: 0490 row: 0 column: 12									
AUTHORS										Seq primer: CACACAGGAACAGCTATGACC									
TITLE										Class: plasmid ends									
										High quality sequence stop: 41.									
										Location/Qualifiers									
										1. .41									
										/organism="Mus musculus"									
										/strain="C57BL/6J"									
										/db_xref="taxon:10090"									
										/clone="UUGC1M0490012"									
										/clone_lib="Mouse 10kb plasmid UUGC1M library"									
										/sex="Male"									
										/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"									
										/note="Vector: PWD42nv; Purified genomic DNA from M.									
										musculus C57BL/6J (male) was obtained from the Jackson									
										Fine Structural analysis of transcription start sites of human									

LOCUS	30 bp DNA	SSS	13-DEC-2000
TA272C090	T. brucei sheared genomic DNA clone 272c09, reverse sequence, genomic survey sequence.		
ACCESSION	AL487818		
VERSION	AL487818.1	GI:11852489	
KEYWORDS	SSS.		
SOURCE	Trypanosoma brucei.		
ORGANISM	Trypanosoma brucei.		
REFERENCE	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.		
AUTHORS	1 (bases 1 to 30) Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk		
COMMENT	Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 Gutat 10.1) was mechanically sheared to give a tight size distribution (to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).		
	Email: nelsayed@tigr.org		
	details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/projects/T_brucei/ .		

FEATURES	source	at http://www.sanger.ac.uk/Projects/I_Lolucei/									
		1. .30									
		/organism="Trypanosoma brucei"									
		/strain="TREU927"									
		/db_xref="taxon:5691"									
		/clone="272c09"									
BASE COUNT		9	a	10	c	3	g	8	t		
ORIGIN											
		Query Match 61.0%; Score 12.2; DB 258; Length 30;									
		Best Local Similarity 82.4%; Pred. No. 5.2e+04;									
		Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;									
Qy	3	taagcagctgccattat 19									
Db	24	TGAGCAGATGCCAGTAT 8									
RESULT	6										
LOCUS	AA789284/c										
DEFINITION	AA789284 40 bp mRNA EST 31-DEC-1998	aj28c05.s1 Soares testis_NHT Homo sapiens cDNA clone 1391624 3'									
		similar to TR:P91373 P91373 SIMILARITY TO RAT TEST PROTEIN. ;, mRNA									
		sequence.									
ACCESSION	AA789284										
VERSION	AA789284.1	GI:2849404									
KEYWORDS	EST.										
SOURCE	human.										
ORGANISM	Homo sapiens										
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;										
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.										
REFERENCE	1 (bases 1 to 40)										
AUTHORS	NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.										
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),										
	Tumor Gene Index										
JOURNAL	Unpublished (1997)										
COMMENT	Contact: Robert Strausberg, Ph.D.										
	Email: cgapbs-r@mail.nih.gov										
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo										
	, Ph.D.										

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio1.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Insert Length: 1216 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES	SOURCE
-----------------	---------------

1. 40
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1391624"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5].
TGTACCAATCTAGTGGAGCGGCCCAATTTTTTTTTTTT 3].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
9 c 12 g t
11 a

	BASE COUNT	ORIGIN
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0
15	0	0
16	0	0
17	0	0
18	0	0
19	0	0
20	0	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	0	0
29	0	0
30	0	0
31	0	0
32	0	0
33	0	0
34	0	0
35	0	0
36	0	0
37	0	0
38	0	0
39	0	0
40	0	0
41	0	0
42	0	0
43	0	0
44	0	0
45	0	0
46	0	0
47	0	0
48	0	0
49	0	0
50	0	0
51	0	0
52	0	0
53	0	0
54	0	0
55	0	0
56	0	0
57	0	0
58	0	0
59	0	0
60	0	0
61	0	0
62	0	0
63	0	0
64	0	0
65	0	0
66	0	0
67	0	0
68	0	0
69	0	0
70	0	0
71	0	0
72	0	0
73	0	0
74	0	0
75	0	0
76	0	0
77	0	0
78	0	0
79	0	0
80	0	0
81	0	0
82	0	0
83	0	0
84	0	0
85	0	0
86	0	0
87	0	0
88	0	0
89	0	0
90	0	0
91	0	0
92	0	0
93	0	0
94	0	0
95	0	0
96	0	0
97	0	0
98	0	0
99	0	0
100	0	0

Query Match 61.08; Score 12.2; DB 12; Length 40;

Best Local Similarity	82.6%;	Pred. NO. 5.5e+04;	
Matches	14;	Conservative	0;
Matches	14;	Mismatches	3;
Matches	14;	Indels	0;
Matches	14;	Gaps	0;

Qy 1 ggtaagcagctgccatt 17
 |||| ||||| |||||
 Db 18 GGTACACAGCTGGCATT 2

RESULT 7

AZ798762	AZ798762	31 bp	DNA	GSS	16-FEB-2001
LOCUS					
DEFINITION	2M0055C18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0055C18 R, DNA sequence.				

ACCESSION AZ798762
 VERSION AZ798762.1 GI:12949191
 KEYWORDS GSS.

KEYWORDS	SOURCE	ORGANISM
<p> Bacterial Fungal Parasitic Protozoan Viral </p>	<p> Bacterial Fungal Parasitic Protozoan Viral </p>	<p> Bacterial Fungal Parasitic Protozoan Viral </p>

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 31)

DUNN, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islan, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Stokes, R., Tingey, A., von Niederhausern, A.,
 and Wright, D. Weiss, R.

TITLE

**JOURNAL
COMMENT**

plasmid inserts
unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 **Std Error:** 0.00
Plate: 0055 **row:** C **column:** 18
Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends
High quality sequence stop: 31.
Location/Qualifiers
1. .31

FEATURES

source

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0055C18"
/clone_lib="Mouse 10kb plasmid UUGC1m library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

8 a 6 c 8 g 9 t
BASE COUNT
ORIGIN

Query Match 59.0%; Score 11.8; DB 249; Length 31;
Best Local Similarity 86.7%; Pred. NO. 8.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 gcagctgccattatt 20
||||| || |||||

Db 6 GCAGCAGCAATTATT 20

RESULT 8

A2311284/c

LOCUS A2311284 33 bp DNA GSS 29-SEP-2000
DEFINITION 1M0026G08R Mouse 10kb plasmid UUGC1m library Mus musculus genomic clone UUGC1M0026G08 R, DNA sequence.

A2311284

A2311284.1 GI:10354094

GSS.

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 33)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,

Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5605

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0026 row: G column: 08

Seq primer: CACACAGGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 33.

Location/Qualifiers

FEATURES

source

1. .33

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0026G08"

/clone_lib="Mouse 10kb plasmid UUGC1m library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

6 a 10 c 9 g 8 t

BASE COUNT

ORIGIN

Query Match 59.0%; Score 11.8; DB 241; Length 33;

Best Local Similarity 86.7%; Pred. NO. 8.5e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggtaagcagctgcc 15

||||| ||||| |||

Db 29 GGTAAGTAGCTCCCA 15

RESULT 9

AA829551/c

LOCUS

AA829551

41 bp mRNA EST 29-APR-1998

DEFINITION

of09h01.s1 NCI-CGAP.C012 Homo sapiens cDNA clone IMAGE:142057 3'

similar to gb:M33552 LYMPHOCYTE-SPECIFIC PROTEIN LSP1 (HUMAN);

mRNA sequence.

AA829551

AA829551.1 GI:2902650

VERSION

EST.

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 41)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Stratagene, Inc.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert Length: 660 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 1.

FEATURES

source
 1. .41
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1420657"
 /clone_lib="NCI_CGAP_Col2"
 /sex="mixed"
 /tissue_type="colon tumor"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI
 ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
 dT. Pooled colon tumors. 5' adaptor sequence: 5'
 GATTCGCACGAG 3' 3' adaptor sequence: 5'
 CTCGATTTTTTTTTTTTTTTT 3' Average insert size: 1.2 kb."

BASE COUNT 10 a 6 c 11 g 14 t
 ORIGIN

Query Match 59.0%; Score 11.8; DB 12; Length 41;
 Best Local Similarity 86.7%; Pred. No. 8.8e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 taagcagctgcatt 17
 ||||| |||||
 Db 19 TAGCAGCCACCAT 5

RESULT 10

TI17561/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

TI17561 45 bp mRNA EST 06-JUN-1994
 mps v160 The blue guys library Saccharomyces cerevisiae cDNA
 sequence upstream of LacZ fusion similar to ORFW, M91073, mRNA
 sequence.
 TI17561 1 GI:458583
 EST.
 Saccharomyces cerevisiae.
 Saccharomyces cerevisiae
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 1 (bases 1 to 45)
 Burns, N., Grimwade, B., Ross-Macdonald, P.B., Choi, E.-Y., Finberg, K.,
 Roeder, G.S. and Snyder, M.
 Large-scale analysis of gene expression, protein localization and
 gene disruption in Saccharomyces cerevisiae
 Genes Dev. 8, 1087-1105 (1994)
 95011603
 Contact: Snyder M
 Department of Biology
 Yale University
 New Haven CT 06520-8103
 Tel: 2034326139
 Fax: 2034326161
 Email: snymicp@yalevm.ycc.yale.edu
 LacZ fusion; Vegetative expression; Beta-gal fusion localization
 pattern;
 50 cytoplasmic spots; Disruption phenotype: inviable, unbudded
 arrest; Fusion: codon12 of ORF next to BRFL. Sequence below near
 or adjacent to lacZ.
 Seq primer: LacZ sequences in transposon.
 Location/Qualifiers

FEATURES

source

1. .45
 /organism="Saccharomyces cerevisiae"
 /db_xref="taxon:4932"
 /clone_lib="The blue guys library"
 /lab_host="E.coli"
 /note="Vector: pRECmtn; A yeast genomic DNA library was
 prepared in the vector pHS56, and subjected to transposon
 mutagenesis with mTn3. This mini-transposon carries lacZ

sequences that lack an initiation codon; expression of
 lacZ is only provided by in frame fusion to yeast coding
 sequence. The yeast genomic DNA carrying the transposon
 was excised from pHS56 and transplaced back onto the yeast
 chromosome. Yeast colonies expressing lacZ were screened
 for in a color assay. A plasmid containing the genomic
 DNA/lacZ fusion junction was recovered from each
 individual yeast colony that expressed lacZ activity.
 These recovered plasmids comprise 'the blue guys library'.
 The fusion junction was then sequenced to identify the
 expressed ORF upstream of the fusion."

BASE COUNT 14 a 13 c 8 g 10 t
 ORIGIN

Query Match 59.0%; Score 11.8; DB 188; Length 45;
 Best Local Similarity 86.7%; Pred. No. 8.9e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 gcagctgcattatt 20
 ||||| |||||
 Db 15 GCAGCTGCCACTAAT 1

RESULT 11

AA662608

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AA662608 50 bp mRNA EST 12-NOV-1997
 nt16a12.s1 NCI_CGAP_Ew1 Homo sapiens cDNA clone IMAGE:1168126
 similar to SW:RL24_HUMAN P38663 60S RIBOSOMAL PROTEIN L24. ;, mRNA
 sequence.
 AA662608
 AA662608.1 GI:2615457
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 50)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
 Ph.D.
 CDNA Library Preparation: David B. Krizman, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40m13 fwd. ET from Amersham.
 Location/Qualifiers

FEATURES

source

1. .50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1168126"
 /clone_lib="NCI_CGAP_Ew1"
 /tissue_type="Ewing's sarcoma"
 /lab_host="DH10B"
 /note="Vector: pAMP10; mRNA made from Ewing's sarcoma,
 cDNA made by oligo-dT priming. Non-directionally cloned.
 Size-selected on agarose gel, average insert size 600 bp.
 Reference: Krizman et al. (1996) Cancer Research
 56:5380-5383."

BASE COUNT 9 a 11 c 14 g 16 t
 ORIGIN

Query Match 59.0%; Score 11.8; DB 10; Length 50;
 Best Local Similarity 86.7%; Pred. No. 9.1e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	3	taagcagctgcatt	17
Db	11	TTAGCAGCAGCCATT	25

RESULT 12

[illegible]

Email: yuzuki@iems.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

Query Match 59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 9.1e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0

RESULT 13

AU103485/C	AU103485	50 bp	EST	05-APR-2001
LOCUS	AU103485	Sugano Homo sapiens	CDNA library	Homo sapiens
DEFINITION	COL04284	CDNA sequence.		
ACCESSION	AU103485			
VERSION	AU103485.1	GI:13553006		
KEYWORDS	EST			
SOURCE	bio			

Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. *Gene* 200 (1-2), 149-156 (1997).

Query Match	59.0%	Score 11.8;	DB 107;	Length 50;
Best Local Similarity	86.7%	Pred. NO. 9.1e+04;		
Matches 13;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

RESULT 14

LOCUS	AU103486	50 bp	mRNA	EST	05-APR-2001
DEFINITION	AU103486	Sugano Homo sapiens	cdna library	Homo sapiens	CDNA clone
ACCESSION	COLFA166	mRNA sequence.			
VERSION	AU103486				
KEYWORDS	AU103486.1	GI:13553007			
SOURCE	EST.				
	human.				

Query Match 59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 9.1e+04;
Matches 13: Conservative 0; Mismatches 2: Indels 0

RESULT 15

AU103487/C

LOCUS AU103487 50 bp mRNA EST 05-APR-2001
DEFINITION AU103487 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
COLF6739, mRNA sequence.
ACCESSION AU103487
VERSION AU103487.1 GI:13553008
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Talra,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.
TITLE Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL Unpublished (2001)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source 1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COLF6739"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 12 a 19 c 11 g 8 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 9.1e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ggtaagcagctgcc 15
||| |
Db 40 GGGAGGCAGCTGCC 26

Search completed: October 2, 2001, 15:00:56
Job time: 10825 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:45 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100b-11

Perfect score: 20

Sequence: 1 agtaccaggtagtcttag 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_cm:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pl1:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rdd:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rdd:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_cm:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vl:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vl:*
- 59: gb_vl2:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rdd:*
- 95: gb_rdd2:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	13.4	67.0	21	10	I30544 Sequence 7
C 2	12.8	64.0	29	10	I75320 Sequence 69
C 3	12.8	64.0	34	9	AR091705 Sequence
C 4	12.8	64.0	48	9	AR075823 Sequence
C 5	12.8	64.0	48	10	E30456 Sequence
C 6	12.4	62.0	18	9	A34802 Method for
C 7	12.4	62.0	22	9	AR067320 Sequence
C 8	12.2	61.0	44	9	AR035220 Sequence

```

9 12.2 61.0 44 9 AR035241 Sequence
10 12 60.0 20 9 AR060552 Sequence
11 12 60.0 39 9 A28774 Oligonucleo
12 11.8 59.0 41 9 AR075947 Sequence
13 11.6 58.0 23 9 A49393 Sequence
14 11.6 58.0 26 9 A97534 Sequence
15 11.6 58.0 26 9 AR013888 Sequence
16 11.6 58.0 26 9 AR033842 Sequence
17 11.6 58.0 26 9 AR042502 Sequence
18 11.6 58.0 26 9 AR058382 Sequence
19 11.6 58.0 26 9 AR088208 Sequence
20 11.6 58.0 26 9 AR089875 Sequence
21 11.6 58.0 35 9 A17379 Nucleotide
22 11.6 58.0 40 9 AR013900 Sequence
23 11.6 58.0 40 9 AR033854 Sequence
24 11.6 58.0 40 9 AR042514 Sequence
25 11.6 58.0 40 9 AR058394 Sequence
26 11.6 58.0 40 9 AR088220 Sequence
27 11.6 58.0 42 9 AR013872 Sequence
28 11.6 58.0 42 9 AR021426 Sequence
29 11.6 58.0 42 9 AR033826 Sequence
30 11.6 58.0 42 9 AR042486 Sequence
31 11.6 58.0 42 9 AR042988 Sequence
32 11.6 58.0 42 9 AR058366 Sequence
33 11.6 58.0 42 9 AR088192 Sequence
34 11.6 58.0 42 10 I43967 Sequence
35 11.6 58.0 42 10 I62979 Sequence
36 11.6 58.0 42 10 I88732 Sequence
37 11.4 57.0 20 10 E28760 Antitumor d
38 11.4 57.0 24 9 A25691 MOUH3FOR ol
39 11.4 57.0 24 9 AR051366 Sequence
40 11.4 57.0 30 10 I36150 Sequence
41 11.4 57.0 38 9 AR061462 Sequence
42 11.4 57.0 38 9 AR108361 Sequence
43 11.4 57.0 38 10 I16318 Sequence
44 11.4 57.0 38 10 I66804 Sequence
45 11.4 57.0 38 10 I84898 Sequence

```

ALIGNMENTS

```

RESULT 1
I30544/c 130544 21 bp DNA PAT 06-FEB-1997
LOCUS Sequence 7 from patent US 5580969.
DEFINITION I30544
ACCESSION I30544
VERSION I30544.1 GI:1821335
KEYWORDS
SOURCE unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 21)
AUTHORS Hoke,G.D., Bradley,M.O., Williams,T.J. and Lee,C.
TITLE Antisense oligonucleotides directed against human ICAM-1 RNA
JOURNAL Patent: US 5580969-A 7 03-DEC-1996;
FEATURES
Location/Qualifiers
1..21
/organism="unknown"
BASE COUNT 5 a 9 c 5 g 2 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 10; Length 21;
Best Local Similarity 93.3%; Pred. No. 6.4e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 gtaccacggtagtc 16
|||
Db 19 GTTCCAGGTGAGTC 5

RESULT 2
I30544/c 130544 21 bp DNA PAT 06-FEB-1997
LOCUS Sequence 7 from patent US 5580969.
DEFINITION I30544
ACCESSION I30544
VERSION I30544.1 GI:1821335
KEYWORDS
SOURCE unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 21)
AUTHORS Hoke,G.D., Bradley,M.O., Williams,T.J. and Lee,C.
TITLE Antisense oligonucleotides directed against human ICAM-1 RNA
JOURNAL Patent: US 5580969-A 7 03-DEC-1996;
FEATURES
Location/Qualifiers
1..21
/organism="unknown"
BASE COUNT 5 a 9 c 5 g 2 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 10; Length 21;
Best Local Similarity 93.3%; Pred. No. 6.4e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 gtaccacggtagtc 16
|||
Db 19 GTTCCAGGTGAGTC 5

RESULT 2
I30544/c 130544 21 bp DNA PAT 06-FEB-1997
LOCUS Sequence 7 from patent US 5580969.
DEFINITION I30544
ACCESSION I30544
VERSION I30544.1 GI:1821335
KEYWORDS
SOURCE unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 21)
AUTHORS Hoke,G.D., Bradley,M.O., Williams,T.J. and Lee,C.
TITLE Antisense oligonucleotides directed against human ICAM-1 RNA
JOURNAL Patent: US 5580969-A 7 03-DEC-1996;
FEATURES
Location/Qualifiers
1..21
/organism="unknown"
BASE COUNT 5 a 9 c 5 g 2 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 10; Length 21;
Best Local Similarity 93.3%; Pred. No. 6.4e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 gtaccacggtagtc 16
|||
Db 19 GTTCCAGGTGAGTC 5

```

```

I75320/c
LOCUS I75320 29 bp DNA PAT 03-APR-1998
DEFINITION Sequence 69 from patent US 5689052.
ACCESSION I75320
VERSION I75320.1 GI:3011461
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 29)
AUTHORS Brown,S.Marie, Dean,D.Allen, Fromm,M.Ernest and Sanders,P.Rigden.
TITLE Synthetic DNA sequences having enhanced expression in
monocotyledonous plants and method for preparation thereof
JOURNAL Patent: US 5689052-A 69 18-NOV-1997;
FEATURES
Location/Qualifiers
1..29
/organism="unknown"
BASE COUNT 6 a 9 c 6 g 8 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 10; Length 29;
Best Local Similarity 87.5%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 agtaccacggtagtc 16
|||
Db 24 AGTGGCCAGGAGAGTC 9

RESULT 3
AR091705 34 bp DNA PAT 07-SEP-2000
LOCUS AR091705
DEFINITION Sequence 15 from patent US 5994505.
ACCESSION AR091705
VERSION AR091705.1 GI:10018459
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 34)
AUTHORS Ting,J.Pan-Yung and Chin,K.
TITLE Forms of class II MHC transactivator (CIITA)
JOURNAL Patent: US 5994505-A 15 30-NOV-1999;
FEATURES
Location/Qualifiers
1..34
/organism="unknown"
BASE COUNT 7 a 11 c 10 g 6 t
ORIGIN
Query Match 64.0%; Score 12.8; DB 9; Length 34;
Best Local Similarity 87.5%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 taccacggtagtc 18
|||
Db 16 TACCCAGCTGTCTCT 31

RESULT 4
AR075823/c 48 bp DNA PAT 30-AUG-2000
LOCUS AR075823
DEFINITION Sequence 6 from patent US 5958700.
ACCESSION AR075823
VERSION AR075823.1 GI:10002569
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 48)
AUTHORS Nadeau,J.G., Pitner,J.Bruce, Linn,C.Preston and Schram,J.L.
TITLE Detection of nucleic acids by fluorescence quenching

```

Renard,A. and Thiry,M. fish Patent: EP 0377349-A 20 11-JUL-1990; EUROGENTEC S.A. Location/Qualifiers 1. .18 /organism="synthetic construct" /db_xref="taxon:32630" 5 a 4 c 7 g 2 t	Query Match 62.0%; Score 12.4; DB 9; Length 18; Best Local Similarity 92.9%; Pred. No. 2.6e+04; Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0
Qy 5 cccaggtgagtctt 18 	
Db 14 CCCAGTCAGTCTT 1	
RESULT 7	
AR067320/c	
LOCUS AR067320 22 bp DNA PAT 29-SEP-1999	
DEFINITION Sequence 668 from patent US 5851760.	
ACCESSION AR067320	
VERSION AR067320.1 GI:5998542	
KEYWORDS	
SOURCE Unknown.	
ORGANISM	
REFERENCE 1 (bases 1 to 22)	
AUTHORS Evans,G.A. and Smith,M.W.	
TITLE Method for generation of sequence sampled maps of complex genomes	
JOURNAL Patent: US 5851760-A 668 22-DEC-1998;	
FEATURES Location/Qualifiers	
source 1.22	
/organism="unknown"	
BASE COUNT 8 a 5 c 4 g 5 t	
ORIGIN	
Query Match 62.0%; Score 12.4; DB 9; Length 22; Best Local Similarity 92.9%; Pred. No. 2.6e+04; Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
Qy 4 acccaggtgagtct 17 	
Db 14 ATCCAGGTGAGTCT 1	
RESULT 8	
AR035220/c	
LOCUS AR035220 44 bp DNA PAT 29-SEP-1999	
DEFINITION Sequence 36 from patent US 5871732.	
ACCESSION AR035220	
VERSION AR035220.1 GI:5951888	
KEYWORDS	
SOURCE Unknown.	
ORGANISM	
REFERENCE 1 (bases 1 to 44)	
AUTHORS Burkly,L.C., Chisholm,P.L., Thomas,D.W., Rosa,M.D. and Rosa,J.J.	
TITLE Anti-CD4 antibody homologs useful in prophylaxis and treatment of AIDS, ARC and HIV infection	
JOURNAL Patent: US 5871732-A 36 16-FEB-1999;	
FEATURES Location/Qualifiers	
source 1..44	
/organism="unknown"	
BASE COUNT 11 a 12 c 14 g 7 t	
ORIGIN	

Query Match 61.0%; Score 12.2; DB 9; Length 44;
 Best Local Similarity 82.4%; Pred. No. 3.2e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 taccacaggtagtcttta 19
 | | | | | | | | | | |
 Db 27 TCCTCAGGTGAGTCCTTA 11

RESULT 9
 AR035241
 LOCUS AR035241 44 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 59 from patent US 5871732.
 ACCESSION AR035241
 VERSION AR035241.1 GI:5951909
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 44)
 AUTHORS Burkly,L.C., Chisholm,P.L., Thomas,D.W., Rosa,M.D. and Rosa,J.J.
 TITLE Anti-CD4 antibody homologs useful in prophylaxis and treatment of
 AIDS, ARC and HIV infection
 JOURNAL Patent: US 5871732-A 59 16-FEB-1999;
 FEATURES Location/Qualifiers
 1..44
 /organism="unknown"
 BASE COUNT 7 a 14 c 12 g 11 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 44;
 Best Local Similarity 82.4%; Pred. No. 3.2e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 taccacaggtagtcttta 19
 | | | | | | | | | | |
 Db 22 TCCTCAGGTGAGTCCTTA 38

RESULT 10
 AR060552
 LOCUS AR060552 20 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 53 from patent US 5840693.
 ACCESSION AR060552
 VERSION AR060552.1 GI:5987002
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 20)
 AUTHORS Eriksson,U., Olofsson,B., Alitalo,K. and Pajusola,K.
 TITLE Vascular endothelial growth factor-B
 JOURNAL Patent: US 5840693-A 53 24-NOV-1998;
 FEATURES Location/Qualifiers
 1..20
 /organism="unknown"
 BASE COUNT 5 a 4 c 6 g 5 t
 ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.5e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 caggtagtcttt 18
 | | | | | | | | | |
 Db 8 CAGGTGAGTCTT 19

RESULT 11
 A28774

LOCUS A28774 39 bp DNA PAT 12-JUN-1995
 DEFINITION Oligonucleotide xol123.
 ACCESSION A28774
 VERSION A28774.1 GI:1248748
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 39)
 AUTHORS
 TITLE PROTEINS PREVENTING INTERACTION BETWEEN A FRAGMENT FC OF AN
 IMMUNOGLOBULIN AND ITS RECEIVER, AND THERAPEUTIC USE, PARTICULARLY
 IN TREATING AFFECTIONS LINKED TO THE HIV VIRUS
 JOURNAL Patent: WO 9108301-A 6 13-JUN-1991;
 FEATURES Location/Qualifiers
 1..39
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 BASE COUNT 12 a 6 c 12 g 9 t
 ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 39;
 Best Local Similarity 100.0%; Pred. No. 4.3e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 agtaccaggtag 12
 | | | | | | | | | |
 Db 15 AGTACCCAGGTG 26

RESULT 12
 AR075947
 LOCUS AR075947 41 bp DNA PAT 30-AUG-2000
 DEFINITION Sequence 8 from patent US 5958713.
 ACCESSION AR075947
 VERSION AR075947.1 GI:10002693
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 41)
 AUTHORS Thastrup,O., Tullin,S.slashed.ren, Poulsen,L.Kongsbak and Bj.o
 slashed.ren,S.petersen.
 TITLE Method of detecting biologically active substances by using green
 fluorescent protein
 JOURNAL Patent: US 5958713-A 8 28-SEP-1999;
 FEATURES Location/Qualifiers
 1..41
 /organism="unknown"
 BASE COUNT 8 a 9 c 12 g 12 t
 ORIGIN

Query Match 59.0%; Score 11.8; DB 9; Length 41;
 Best Local Similarity 86.7%; Pred. No. 5.7e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ccaggtgagtccttag 20
 | | | | | | | | | |
 Db 20 CCGGTGAGTCATAG 34

RESULT 13
 A49393/c
 LOCUS A49393 23 bp DNA PAT 07-MAR-1997
 DEFINITION Sequence 11 from Patent WO9607740.
 ACCESSION A49393
 VERSION A49393.1 GI:2302870
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:00:56 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-11

Perfect score: 20

Sequence: 1 agtaccagggtgagtcttag 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_est4.*
- 5: gb_est5.*
- 6: gb_est6.*
- 7: gb_est7.*
- 8: gb_est8.*
- 9: gb_est9.*
- 10: gb_est10.*
- 11: gb_est11.*
- 12: gb_est12.*
- 13: gb_est13.*
- 14: gb_est14.*
- 15: gb_est15.*
- 16: gb_est16.*
- 17: gb_est17.*
- 18: gb_est18.*
- 19: gb_est19.*
- 20: gb_est20.*
- 21: gb_est21.*
- 22: gb_est22.*
- 23: gb_est23.*
- 24: gb_est24.*
- 25: gb_est25.*
- 26: gb_est26.*
- 27: gb_est27.*
- 28: gb_est28.*
- 29: gb_est29.*
- 30: gb_est30.*
- 31: gb_est31.*
- 32: gb_est32.*
- 33: gb_est33.*
- 34: gb_est34.*
- 35: gb_est35.*
- 36: gb_est36.*
- 37: gb_est37.*
- 38: gb_est38.*
- 39: gb_est39.*
- 40: gb_est40.*
- 41: gb_est41.*
- 42: gb_est42.*
- 43: gb_est43.*
- 44: gb_est44.*
- 45: gb_est45.*
- 46: gb_est46.*
- 47: gb_est47.*

- 44: em_esthum10.*
- 45: em_esthum11.*
- 46: em_esthum12.*
- 47: em_esthum13.*
- 48: em_esthum14.*
- 49: em_esthum15.*
- 50: em_esthum16.*
- 51: em_esthum17.*
- 52: em_esthum18.*
- 53: em_esthum19.*
- 54: em_esthum20.*
- 55: em_esthum21.*
- 56: em_esthum22.*
- 57: em_esthum23.*
- 58: em_esthum24.*
- 59: em_esthum25.*
- 60: em_esthum26.*
- 61: em_esthum27.*
- 62: em_esthum28.*
- 63: em_estin1.*
- 64: em_estin2.*
- 65: em_estin3.*
- 66: em_estin4.*
- 67: em_estin5.*
- 68: em_estom1.*
- 69: em_estom2.*
- 70: em_estov1.*
- 71: em_estov2.*
- 72: em_estpl1.*
- 73: em_estpl2.*
- 74: em_estpl3.*
- 75: em_estpl4.*
- 76: em_estpl5.*
- 77: em_estpl6.*
- 78: em_estpl7.*
- 79: em_estpl8.*
- 80: em_estpl9.*
- 81: em_estpl10.*
- 82: em_estro1.*
- 83: em_estro2.*
- 84: em_estro3.*
- 85: em_estro4.*
- 86: em_estro5.*
- 87: em_estro6.*
- 88: em_estro7.*
- 89: em_estro8.*
- 90: em_estro9.*
- 91: em_estro10.*
- 92: em_estro11.*
- 93: em_estro12.*
- 94: em_estro13.*
- 95: em_estro14.*
- 96: em_estro15.*
- 97: em_estro16.*
- 98: em_estro17.*
- 99: em_estro18.*
- 100: em_estro19.*
- 101: em_estro20.*
- 102: gb_est25.*
- 103: gb_est26.*
- 104: gb_est27.*
- 105: gb_est28.*
- 106: gb_est29.*
- 107: gb_est30.*
- 108: gb_est31.*
- 109: gb_est32.*
- 110: gb_est33.*
- 111: gb_est34.*
- 112: gb_est35.*
- 113: gb_est36.*
- 114: gb_est37.*
- 115: gb_est38.*
- 116: gb_est39.*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: em_est100:*
170: em_est101:*
171: em_est102:*
172: em_est103:*
173: em_est104:*
174: em_est105:*
175: em_est106:*
176: em_est107:*
177: em_est108:*
178: em_est109:*
179: em_est110:*
180: em_est111:*
181: em_est112:*
182: em_est113:*
183: em_est114:*
184: em_est115:*
185: em_est116:*
186: em_est117:*
187: em_est118:*
188: em_est119:*
189: em_est120:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0347 row: C column: 11
Seq primer: CGTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 25.
FEATURES
Location/Qualifiers
source
1. .25
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0347C11"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydronamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gil4732114/gb/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

Query Match 64.0%; Score 12.8; DB 245; Length 25;
Best Local Similarity 87.5%; Pred. No. 1.8e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 ccaggtgagcttag 20
1 | | | | | | | | | |
Db 5 CTCAGGTGAGTCTTTG 20

RESULT 2
AZ666528
LOCUS 37 bp DNA GSS 14-DEC-2000
DEFINITION IM0548A18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0548A18 R, DNA sequence.
ACCESSION AZ666528
VERSION AZ666528.1 GI:11803674
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 37)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D.,Weiss,R.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	12.8	64.0	25	245	AZ506197	IM0347C11
2	12.6	63.0	37	247	AZ666528	IM0548A18
c 3	12.4	62.0	32	241	AZ303920	IM0003F19
4	12	60.0	27	159	L32037	HUMXPFI11B
5	11.8	59.0	50	107	AU105613	AU105613
c 6	11.6	58.0	46	250	AZ826374	AZ826374
c 7	11.6	58.0	47	156	D38689	HUMCI264.Hu
8	11.2	56.0	31	249	AZ774468	AZ774468
c 9	11.2	56.0	33	249	AZ799299	2M0003017
c 10	11.2	56.0	34	249	AZ763845	2M0056M03
c 11	11.2	56.0	43	244	AZ480568	1M0559C04
c 12	11.2	56.0	43	244	AZ480568	AZ480568
c 13	11.2	56.0	43	249	AZ806794	AZ806794
c 14	11.2	56.0	48	258	TA117F12P	2M0069006
15	11	55.0	50	107	AU106341	AL462744 T. brucei
c 16	11	55.0	34	188	R84653	AU106341
c 17	11	55.0	48	18	AI318658	Y038C05.r1
c 18	11	55.0	50	107	AU102283	AI318658 Ta80A03.x
19	11	55.0	50	107	AU102992	AU102283 AU102283
c 20	10.8	54.0	27	244	AZ495501	AU105601
21	10.8	54.0	28	241	AZ313467	AZ495501 IM0331P21
c 22	10.8	54.0	33	242	AZ382781	AZ313467 IM0029P06
c 23	10.8	54.0	38	245	AZ508030	AZ382781 IM0140N08
c 24	10.8	54.0	40	258	TA69H03Q	AZ508030 IM0349P22
c 25	10.8	54.0	42	243	AZ427649	AL458516 T. brucei
c 26	10.8	54.0	43	249	AZ785625	AZ427649 IM0209A15
27	10.8	54.0	45	21	AI538057	AZ785625 2M0029J08
c 28	10.8	54.0	50	107	AU105612	AI538057 to83B07.x
c 29	10.6	53.0	31	249	AZ770290	AU105612
c 30	10.6	53.0	37	102	AI862676	AZ770290 IM0571G18
31	10.6	53.0	37	247	AZ644324	AI862676 wj27f04.x
c 32	10.6	53.0	42	11	AA741237	AZ644324 IM0508A13
c 33	10.6	53.0	43	103	AI917489	AA741237 ny96B05.s
34	10.6	53.0	43	187	S0655	AI917489 ts84f01.x
35	10.6	53.0	46	159	H98043	S0655 yj57g03.r1
36	10.6	53.0	46	249	AZ764208	H98043 vw10f05.r1
37	10.6	53.0	48	250	AZ819513	AZ764208 IM0560101
38	10.6	53.0	49	168	BF740213	AZ819513 2M0091L11
39	10.6	53.0	49	189	WI2708	BF740213 hu57b11.x
40	10.6	53.0	50	107	AU106954	WI2708 ma52C01.r1
c 41	10.4	52.0	23	243	AZ433756	AU106954 AU106954
c 42	10.4	52.0	24	244	AZ478673	AZ433756 IM0219J18
c 43	10.4	52.0	25	191	W95789	AZ478673 IM0298J20
44	10.4	52.0	26	250	AZ827167	W95789 ze07g08.r1
c 45	10.4	52.0	26	258	TA30H08Q	AZ827167 2M0103N22
						AL452912 T. brucei

ALIGNMENTS

RESULT 1
AZ506197
LOCUS 25 bp DNA GSS 05-OCT-2000
DEFINITION IM0347C11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0347C11 F, DNA sequence.
ACCESSION AZ506197
VERSION AZ506197.1 GI:10687513
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 25)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0548 row: A column: 18
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 37.

FEATURES

source
1. 37
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M003F19"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 14 a 7 c 8 g 8 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 247; Length 37;
Best Local Similarity 78.9%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 agtaccagggtgagctctta 19
||| ||||| ||||| |||||
Db 13 AGAACCCATGTGTAATGTTA 31

RESULT 3
L32037 LOCUS AZ303920 32 bp DNA GSS 29-SEP-2000
DEFINITION lM0003F19R Mouse 10kb plasmid UUCG1M library Mus musculus genomic clone UUCG1M0003F19 R, DNA sequence.

ACCESSION AZ303920
VERSION AZ303920.1 GI:10339372
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 32)
REFERENCE Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT
Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0003 row: F column: 19
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 32.

FEATURES

source
1. 32
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG1M0003F19"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 5 a 10 c 9 g 8 t
ORIGIN

Query Match 62.0%; Score 12.4; DB 241; Length 32;
Best Local Similarity 92.9%; Pred. No. 3e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 agtaccagggtgag 14
||| ||||| |||||
Db 31 AGGACCCAGGTGAG 18

RESULT 4
L32037 LOCUS HUMXP1F11B Human placenta Homo sapiens cDNA clone XPG66B, mRNA 02-AUG-1995
DEFINITION L32037 HUMXP1F11B Human placenta Homo sapiens cDNA clone XPG66B, mRNA sequence.

ACCESSION L32037
VERSION L32037.1 GI:927081
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 27)
REFERENCE Lee, C.-C., Yazdani, A., Wehnert, M., Bailey, J., Couch, L., Xiong, M.,

Coolbaugh, M.I., Chinault, C.A., Baldini, A., Lindsay, E.A., Zhao, Z.-Y.
and Caskey, C.T.H.
Isolation of chromosome-specific genes by reciprocal probing of
arrayed cDNAs and cosmid libraries
Hum. Mol. Genet. 4, 1373-1380 (1995)
96090257
Contact: Caskey, C.T.H.

TITLE

Isolation of chromosome-specific genes by reciprocal probing of
arrayed cDNAs and cosmid libraries

JOURNAL
MEDLINE
COMMENT

FEATURES

source

1. .27
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Xq27.3-q28"
/clone="XP66G8"
/note="Arrayed cDNAs and cosmid libraries from human
placental tissue"

BASE COUNT 7 a 4 c 7 g 9 t

ORIGIN

Query Match 60.0%; Score 12; DB 159; Length 27;
Best Local Similarity 75.0%; Pred. No. 4.8e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 agtaccaggtagctttag 20

||||| ||| ||| |||

Db 5 AATACCTGATGCGCTTTAG 24

RESULT 5

AU105613

LOCUS

DEFINITION AU105613 50 bp mRNA EST 05-APR-2001
Sugano Homo sapiens cDNA library Homo sapiens cDNA clone

HP22150, mRNA sequence.

ACCESSION AU105613

VERSION AU105613.1 GI:13555134

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata

, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo

, K., Suyama, A. and Sugano, S.

Fine Structural analysis of transcription start sites of human

mRNAs using full-length enriched and 5'-end enriched cDNA libraries

Unpublished (2001)

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano

, S. Construction and characterization of a full length-enriched and

a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

1. .50
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEP22150"
/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 6 a 7 c 22 g 15 t

ORIGIN

Query Match 59.0%; Score 11.8; DB 107; Length 50;

Best Local Similarity 86.7%; Pred. No. 6.5e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 ccaggtagctttag 20

||||| ||| |||

Db 13 CCGGGTGAAGCTTAG 27

RESULT 6

AZ826374/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0102 row: K column: 06

Seq primer: CGTGTAAACGACGGCCAGT

Class: plasmid ends

High quality sequence stop: 46.

Location/Qualifiers

1. .46

source

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0102K06"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (GI4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 17 a 14 c 9 g 6 t

ORIGIN

Query Match 58.0%; Score 11.6; DB 250; Length 46;

Best Local Similarity 77.8%; Pred. No. 8.2e+04;

Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 agtaccaggtagctt 18

||||| ||| ||| |||

Db 13 CCGGGTGAAGCTTAG 27

```

Db 21 AGTATCAGGTGTGCTT 4

RESULT 7
LOCUS D38689 47 bp mRNA EST 31-MAY-1995
DEFINITION HUMC1264 Human chromosome 8 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION D38689
VERSION D38689.1 GI:807747
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 47)
AUTHORS Kovama,K., Sufo,K. and Nakamura,Y.
TITLE Isolation of 115 human chromosome 8-specific expressed-sequence
tags by exon amplification
JOURNAL Genomics 26, 245-253 (1995)
MEDLINE 95324915
COMMENT Contact: Yusuke Nakamura
Institute of Medical Science
University of Tokyo
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan
Tel: 81-3-5449-5372
Fax: 81-3-5449-5433
Email: yusuke@ims.u-tokyo.ac.jp.
Location/Qualifiers
1. 47
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="8"
/clone.lib="Human chromosome 8"
/note="Exon-like sequence on chromosome 8; exon-trapping
method"

BASE COUNT 21 a 9 c 9 g 8 t
ORIGIN

Query Match 58.0%; Score 11.6; DB 156; Length 47;
Best Local Similarity 77.8%; Pred. No. 8.2e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 taccacagtgagctcttag 20
||||||| | |||||
Db 21 TACCCAGGTAATCTTCG 4

RESULT 8
LOCUS AZ774468 31 bp DNA GSS 16-FEB-2001
DEFINITION 2M0003017R Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC2M0003017 R, DNA sequence.
ACCESSION AZ774468
VERSION AZ774468.1 GI:12899949
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 31)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0003 row: 0 column: 17
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 31.
Location/Qualifiers
1. 31
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC2M0003017"
/clone.lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 9 a 6 c 10 t
ORIGIN

Query Match 56.0%; Score 11.2; DB 249; Length 31;
Best Local Similarity 81.2%; Pred. No. 1.3e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 acccagtgagctcta 19
||||||| | |||||
Db 10 ACCCAGCTGGGTTTGA 25

RESULT 9
LOCUS AZ799299 33 bp DNA GSS 16-FEB-2001
DEFINITION 2M0056M03R Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC2M0056M03 R, DNA sequence.
ACCESSION AZ799299
VERSION AZ799299.1 GI:12950275
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 33)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

```

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0056 row: M column: 03

Seq primer: CACACAGGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 33.

Location/Qualifiers

1. .33

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0056M03"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/notes="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gll14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

11 a 9 c 6 g 7 t

BASE COUNT
ORIGIN

Query Match 56.0%; Score 11.2; DB 249; Length 33;

Best Local Similarity 81.2%; Pred. No. 1.3e+05;

Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 agtaccacagtgatgc 16

||| |||| |||||

Db 33 AGTACCAGTTGAGTC 18

RESULT 10
AZ763845/c

LOCUS AZ763845 34 bp DNA GSS 16-FEB-2001

DEFINITION IM0559C04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0559C04 R, DNA sequence.

ACCESSION AZ763845

VERSION AZ763845.1 GI:12875290

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 34)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

JOURNAL

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0559 row: C column: 04

Seq primer: CACACAGGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 34.

Location/Qualifiers

1. .34

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0559C04"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/notes="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gll14732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

9 a 12 c 9 g 4 t

BASE COUNT
ORIGIN

Query Match 56.0%; Score 11.2; DB 249; Length 34;

Best Local Similarity 81.2%; Pred. No. 1.3e+05;

Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 agtaccacagtgatgc 16

||| ||||| |||||

Db 28 AGTCCCCAGGTGGCC 13

RESULT 11
AZ480568

LOCUS AZ480568 43 bp DNA GSS 04-OCT-2000

DEFINITION IM0302M10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0302M10 F, DNA sequence.

ACCESSION AZ480568

VERSION AZ480568.1 GI:10641633

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 43)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

JOURNAL

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0302 row: M column: 10
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 43.

FEATURES

source

```

1. .43
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUC1M0302M10"
  /clone_lib="Mouse 10kb plasmid UUC1M library"
  /sex="Male"
  /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
  /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      11 a      7 c      17 g      8 t
ORIGIN

```

Query Match 56.0%; Score 11.2; DB 244; Length 43;
Best Local Similarity 81.2%; Pred. NO. 1.3e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY  2 gtaccaggtgagctct 17
    ||| |||| |||||
DB   18 GTAGACAGGAGAGTCT 33

```

```

RESULT 12
LOCUS   A2806794          43 bp      DNA          GSS          20-FEB-2001
DEFINITION  2M0069006F Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC2M0069006 F, DNA sequence.
ACCESSION  A2806794
VERSION    A2806794.1 GI:12970500
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 43)
AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE     Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL   Unpublished (2000)
COMMENT   Contact: Robert B. Weiss

```

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0069 row: O column: 06
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 43.

FEATURES

source

```

1. .43
  /organism="Mus musculus"
  /strain="C57BL/6J"
  /db_xref="taxon:10090"
  /clone="UUC2M0069006"
  /clone_lib="Mouse 10kb plasmid UUC1M library"
  /sex="Male"
  /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
  /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      8 a      5 c      9 g      21 t
ORIGIN

```

Query Match 56.0%; Score 11.2; DB 249; Length 43;
Best Local Similarity 81.2%; Pred. NO. 1.3e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY  4 acccaggtgagctctta 19
    ||| ||| |||||
DB   34 ACCAAGGCAAGTCTTA 19

```

```

RESULT 13
LOCUS   TAIL7F12P/c      48 bp      DNA          GSS          13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 117f12, forward sequence, genomic survey sequence.
ACCESSION  AL462744
VERSION    AL462744.1 GI:11832461
KEYWORDS   GSS.
SOURCE     Trypanosoma brucei.
ORGANISM   Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma. 1 (bases 1 to 48)
AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R., Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE     Direct Submission
JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and

```


nhlsanger.ac.uk
 Constructed at the Institute for Genomic Research (TIGR),
 Rockville, MD. Genomic DNA isolated from a cloned population of
 Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
 to give a tight size distribution (4 kb). The v + i method used for the library construction is
 described in detail in Smith, H. and Venter, J.C. (Making small
 insert libraries for whole genome shotgun sequencing projects. In
 Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Barrell, Oxford University Press, 1999).
 Email: nelsayed@tigr.org
 Details of T. brucei sequencing at the Sanger Centre are available
 at <http://www.sanger.ac.uk/Projects/T-brucei/>.

FEATURES

source

1. .48
 /organism="Trypanosoma brucei"
 /strain="TREU927"
 /db_xref="taxon:5691"
 /clone="117f12"

BASE COUNT 12 a 10 c 15 g 11 t

ORIGIN

Query Match 56.0%; Score 11.2; DB 258; Length 48;
 Best Local Similarity 81.2%; Pred. No. 1.4e+05;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 gtaccagggtgagctt 17

Db 16 GTACCCACGTCGCT 1

RESULT 14

AU106341

LOCUS

DEFINITION AU106341 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 COL03684, mRNA sequence.

ACCESSION AU106341

VERSION AU106341.1

KEYWORDS GI:13555862

EST.

SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 50)
 Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 H., Ota,T., Isogai,T., Tanaka,T., Nakamura,T., Morishita,S., Okubo
 K., Suyama,A. and Sugano,S.

1. Fine Structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries
 Unpublished (2001)

JOURNAL

COMMENT

Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

1. .50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="COL03684"
 /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT

7 a 19 c 16 g 8 t

ORIGIN

Query Match 56.0%; Score 11.2; DB 107; Length 50;
 Best Local Similarity 81.2%; Pred. No. 1.4e+05;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 gtaccagggtgagctt 17

Db 11 GTACCGCGCTGAGCT 26

RESULT 15

R84653

LOCUS

DEFINITION

R84653
 Y038c05.r1 Soares adult brain N2b4HB55Y Homo sapiens cDNA clone
 IMAGE:180200 5' similar to gb:M92302 DIHYDROXYRINDINE-SENSITIVE
 L-TYPE, BRAIN CALCIUM CHANNEL (HUMAN);, mRNA sequence.

ACCESSION R84653

VERSION R84653.1

KEYWORDS GI:943059

EST.

SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 34)
 Hillier,L., Clark,N., Dubuque,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
 M., Hultman,M., Kucaba,T., Soares,M., Tan,F., Trevaskis,E., Waterston
 Rifkin,L., Rohlfing,T., Soares,M., Wilson,R.

TITLE

JOURNAL

COMMENT

The WashU-Merck EST Project
 Unpublished (1995)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 High quality sequence starts: 1
 High quality sequence stops: 1
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Seq primer: M13RPI
 High quality sequence stop: 1.

FEATURES

source

1. .34
 /organism="Homo sapiens"
 /db_xref="GDB:3826844"
 /db_xref="taxon:9606"
 /clone="IMAGE:180200"
 /clone_lib="Soares adult brain N2b4HB55Y"
 /sex="Male"
 /dev_stage="55-year old"
 /lab_host="DH10B (ampicillin resistant)"
 /notes="Organ: brain; Vector: pT7T3D (Pharmacia) with a
 modified polylinker; Site:1: Not 1: Site 2: Eco RI: 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAGTGGAGCGCGCTTTTTTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 53. Library constructed by Bento
 Soares and M.Fatima Bonaldo. The adult brain RNA was
 provided by Dr. Donald H. Gilden. Tissue was acquired
 17-18 hours after death which occurred in consequence of a
 ruptured aortic aneurysm. RNA was prepared from a pool of
 tissues representing the following areas of the brain:
 frontal, parietal, temporal and occipital cortex from the
 left and right hemispheres, subcortical white matter,
 basal ganglia, thalamus, cerebellum, midbrain, pons and
 medulla."

BASE COUNT

11 a 11 c 6 g 5 t 1 others

ORIGIN

Query Match 55.0%; Score 11; DB 188; Length 34;
 Best Local Similarity 70.0%; Pred. No. 1.7e+05;

	Matches	14;	Conservative	0;	Mismatches	6;	Indels	0;	Gaps	0;
Qy	1	agtaccaggtagtcttag	20							
Db	4	AGTCCCGAGGAAGNCTCAG	23							

Search completed: October 2, 2001, 15:00:58
Job time: 10827 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 2, 2001, 15:56:45 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-12
Perfect score: 20
Sequence: 1 cctgacatcagtagcatctc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_om:*

8: gb_ov:*

9: gb_pat1:*

10: gb_pat2:*

11: gb_ph:*

12: gb_pl1:*

13: gb_pl2:*

14: gb_pl3:*

15: gb_pl4:*

16: em_ba1:*

17: em_ba2:*

18: em_fun:*

19: em_htgo_hum:*

20: em_htgo_inv:*

21: em_htgo_rod:*

22: em_htg_hum1:*

23: em_htg_hum2:*

24: em_htg_hum3:*

25: em_htg_hum4:*

26: em_htg_hum5:*

27: em_htg_hum6:*

28: em_htg_hum7:*

29: em_htg_hum8:*

30: em_htg_inv1:*

31: em_htg_inv2:*

32: em_htg_other:*

33: em_htg_rod:*

34: em_hum1:*

35: em_hum2:*

36: em_hum3:*

37: em_hum4:*

38: em_hum5:*

39: em_hum6:*

40: em_hum7:*

41: em_in:*

42: em_om:*

43: em_or:*

44: em_ov:*

45: em_pat:*

46: em_ph:*

47: em_pl:*

48: em_ro:*

49: em_sts:*

50: em_sy:*

51: em_un:*

52: em_vi:*

53: gb_sts1:*

54: gb_sts2:*

55: gb_sts3:*

56: gb_sy:*

57: gb_un:*

58: gb_vil:*

59: gb_vil2:*

60: gb_htg1:*

61: gb_htg2:*

62: gb_htg3:*

63: gb_htg4:*

64: gb_htg5:*

65: gb_htg6:*

66: gb_htg7:*

67: gb_htg8:*

68: gb_htg9:*

69: gb_htg10:*

70: gb_htg11:*

71: gb_htg12:*

72: gb_htg13:*

73: gb_htg14:*

74: gb_htg15:*

75: gb_htg16:*

76: gb_htg17:*

77: gb_htg18:*

78: gb_htg19:*

79: gb_htg20:*

80: gb_htg21:*

81: gb_htg22:*

82: gb_htg23:*

83: gb_htg24:*

84: gb_htg25:*

85: gb_pr1:*

86: gb_pr2:*

87: gb_pr3:*

88: gb_pr4:*

89: gb_pr5:*

90: gb_pr6:*

91: gb_pr7:*

92: gb_pr8:*

93: gb_pr9:*

94: gb_rol:*

95: gb_ro2:*

96: gb_in4:*

97: gb_pr10:*

98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	13.8	69.0	17	9 AR046638	AR046638 Sequence
C 2	13.8	69.0	17	10 I53690	I53690 Sequence 14
C 3	13.8	69.0	21	10 AX095079	AX095079 Sequence
C 4	13.8	69.0	30	10 I27043	I27043 Sequence 54
C 5	13.6	68.0	42	9 AR021414	AR021414 Sequence
C 6	13.6	68.0	42	9 AR042976	AR042976 Sequence
C 7	13.6	68.0	42	10 I43923	I43923 Sequence 12
C 8	13.6	68.0	42	10 I62967	I62967 Sequence 64

```

c 9 13.6 68.0 42 10 188720 188720 Sequence 64
c 10 13.6 68.0 45 97 HUMIGHABH L06994 Human Ig re
c 11 13.2 66.0 30 9 AR063732 Sequence 17
c 12 13.2 66.0 30 10 I30095 I30095 Sequence 17
c 13 13.2 66.0 23 9 AR020665 AR020665 Sequence 18
c 14 12.8 64.0 21 9 A51533 A51533 Sequence 18
c 15 12.8 64.0 21 9 A51561 A51561 Sequence 21
c 16 12.8 64.0 21 9 A87815 A87815 Sequence 11
c 17 12.8 64.0 21 9 A91147 A91147 Sequence 11
c 18 12.8 64.0 21 9 AR084287 AR084287 Sequence 11
c 19 12.8 64.0 21 10 I95565 I95565 Sequence 20
c 20 12.8 64.0 27 10 I25860 I25860 Sequence 20
c 21 12.8 64.0 27 10 I43497 I43497 Sequence 20
c 22 12.4 62.0 18 9 AR047426 AR047426 Sequence 11
c 23 12.4 62.0 18 10 I54478 I54478 Sequence 22
c 24 12.4 62.0 38 9 AR047211 AR047211 Sequence 22
c 25 12.4 62.0 38 9 AR057887 AR057887 Sequence 20
c 26 12.4 62.0 38 10 I54263 I54263 Sequence 20
c 27 12.4 62.0 41 9 AR050319 AR050319 Sequence 20
c 28 12.2 61.0 27 9 A22449 A22449 JH reporter
c 29 12.2 61.0 31 9 AR091207 AR091207 Sequence 17
c 30 12.2 61.0 40 9 AX036346 AX036346 Sequence 3
c 31 12.2 61.0 42 9 AR021372 AR021372 Sequence 3
c 32 12.2 61.0 42 9 AR042934 AR042934 Sequence 3
c 33 12.2 61.0 42 10 I62925 I62925 Sequence 20
c 34 12.2 61.0 42 10 I88678 I88678 Sequence 20
c 35 12.2 61.0 44 9 AR077278 AR077278 Sequence 3
c 36 12.2 61.0 48 10 I05886 I05886 Sequence 3
c 37 12.2 61.0 48 10 I05931 I05931 Sequence 3
c 38 12.2 61.0 48 10 I09165 I09165 Sequence 3
c 39 12.2 60.0 20 9 AR073966 AR073966 Sequence 35
c 40 12.2 60.0 20 9 AR010521 AR010521 Sequence 35
c 41 12.2 60.0 20 10 I27265 I27265 Sequence 1/19
c 42 12.2 60.0 24 9 A09906 A09906 Probe. 1/19
c 43 12.2 60.0 29 9 AX002747 AX002747 Sequence
c 44 12.2 60.0 29 9 AX003151 AX003151 Sequence
c 45 12.2 60.0 29 9 AX018567 AX018567 Sequence

ALIGNMENTS

RESULT 1
AR046638/c 17 bp DNA PAT 29-SEP-1999
LOCUS AR046638 Sequence 1431 from patent US 5817796.
DEFINITION AR046638
ACCESSION AR046638
VERSION AR046638.1 GI:5968103
KEYWORDS
SOURCE
ORGANISM
REFERENCE
Unclassified.
1 (bases 1 to 17)
AUTHORS Stinchcomb,D.T., Draper,K., McSwiggen,J. and Jarvis,T.
TITLE C-myb ribozymes having 2'-5'-linked adenylyate residues
JOURNAL Patent: US 5817796-A 1431 06-OCT-1998;
FEATURES
Location/Qualifiers
1..17
/organism="unknown"
BASE COUNT 2 a 4 c 5 g 6 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 17;
Best Local Similarity 88.2%; Pred. No. 9.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 gacatcagtagcatctc 20
Db 17 GACATCAGGAGCAACTC 1

RESULT 2
AR046638/c 17 bp DNA PAT 07-OCT-1997
LOCUS AR046638 Sequence 1431 from patent US 5646042.
DEFINITION AR046638
ACCESSION AR046638
VERSION AR046638.1 GI:2474893
KEYWORDS
SOURCE
ORGANISM
REFERENCE
Unclassified.
1 (bases 1 to 17)
AUTHORS Stinchcomb,D.T., Draper,K., McSwiggen,J. and Jarvis,T.
TITLE C-myb targeted ribozymes
JOURNAL Patent: US 5646042-A 1431 08-JUL-1997;
FEATURES
Location/Qualifiers
1..17
/organism="unknown"
BASE COUNT 2 a 4 c 5 g 6 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 17;
Best Local Similarity 88.2%; Pred. No. 9.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 gacatcagtagcatctc 20
Db 17 GACATCAGGAGCAACTC 1

RESULT 3
AX095079/c 21 bp DNA PAT 30-MAR-2001
LOCUS AX095079 Sequence 257 from Patent WO0118250.
DEFINITION AX095079
ACCESSION AX095079
VERSION AX095079.1 GI:13511282
KEYWORDS
SOURCE
ORGANISM
REFERENCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 21)
AUTHORS Lander,E.S., Gargill,M., Ireland,J.S., Bolk,S., Daley,G.Q. and
McCarthy,J.J.
TITLE Single nucleotide polymorphisms in genes
JOURNAL Patent: WO 0118250-A 257 15-MAR-2001;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
1..21
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 7 a 4 c 6 g 3 t 1 others
ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 21;
Best Local Similarity 78.9%; Pred. No. 9.3e+03;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 cctgacatcagtagcatctc 19
Db 19 CTTGCCATYAGTGGCATCT 1

RESULT 4
I27043/c 30 bp DNA PAT 06-FEB-1997
LOCUS I27043 Sequence 64 from patent US 5563036.
DEFINITION I27043
ACCESSION I27043
VERSION I27043.1 GI:1817819
KEYWORDS
SOURCE
ORGANISM

```

Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Peterson,M.G., Baichwal,V.R. and Strulovici,B.
TITLE Transcription factor-DNA binding assay
JOURNAL Patent: US 5563036-A 64 08-OCT-1996;
FEATURES Location/Qualifiers
source
1..30
BASE COUNT 8 a 4 c 8 g 10 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 30;
Best Local Similarity 88.2%; Pred. No. 9.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 cctgacatcagtagcat 17
||||| ||||| |||||
Db 27 CCTGATCATTAGCAT 11

RESULT 5
LOCUS AR021414 42 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 64 from patent US 5789650.
ACCESSION AR021414
VERSION AR021414.1 GI:3976029
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 42)
AUTHORS Lonberg,N. and Kay,R.M.
TITLE Transgenic non-human animals for producing heterologous antibodies
JOURNAL Patent: US 5789650-A 64 04-AUG-1998;
FEATURES Location/Qualifiers
source
1..42
BASE COUNT 5 a 8 c 15 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 42;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cctgacatcagtagcatctc 20
||||| ||||| |||||
Db 35 CCAGACATCAAAAGCATCAC 16

RESULT 6
LOCUS AR042976 42 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 64 from patent US 5814318.
ACCESSION AR042976
VERSION AR042976.1 GI:5963984
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 42)
AUTHORS Lonberg,N. and Kay,R.M.
TITLE Transgenic non-human animals for producing heterologous antibodies
JOURNAL Patent: US 5814318-A 64 29-SEP-1998;
FEATURES Location/Qualifiers
source
1..42
BASE COUNT 5 a 8 c 15 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 42;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cctgacatcagtagcatctc 20
||||| ||||| |||||
Db 35 CCAGACATCAAAAGCATCAC 16

Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cctgacatcagtagcatctc 20
||||| ||||| |||||
Db 35 CCAGACATCAAAAGCATCAC 16

RESULT 7
LOCUS I43923 42 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 12 from patent US 5633425.
ACCESSION I43923
VERSION I43923.1 GI:2469021
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 42)
AUTHORS Lonberg,N. and Kay,R.M.
TITLE Transgenic non-human animals capable of producing heterologous antibodies
JOURNAL Patent: US 5633425-A 12 27-MAY-1997;
FEATURES Location/Qualifiers
source
1..42
BASE COUNT 5 a 8 c 15 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 42;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cctgacatcagtagcatctc 20
||||| ||||| |||||
Db 35 CCAGACATCAAAAGCATCAC 16

RESULT 8
LOCUS I62967 42 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 64 from patent US 5661016.
ACCESSION I62967
VERSION I62967.1 GI:2480675
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 42)
AUTHORS Lonberg,N. and Kay,R.M.
TITLE Transgenic non-human animals capable of producing heterologous antibodies of various isotopes
JOURNAL Patent: US 5661016-A 64 26-AUG-1997;
FEATURES Location/Qualifiers
source
1..42
BASE COUNT 5 a 8 c 15 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 42;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cctgacatcagtagcatctc 20
||||| ||||| |||||
Db 35 CCAGACATCAAAAGCATCAC 16

RESULT 9
LOCUS I88720 42 bp DNA PAT 10-AUG-1998
DEFINITION Sequence 64 from patent US 5661016.
ACCESSION I88720
VERSION I88720.1 GI:2480675
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 42)
AUTHORS Lonberg,N. and Kay,R.M.
TITLE Transgenic non-human animals capable of producing heterologous antibodies of various isotopes
JOURNAL Patent: US 5661016-A 64 26-AUG-1997;
FEATURES Location/Qualifiers
source
1..42
BASE COUNT 5 a 8 c 15 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 42;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cctgacatcagtagcatctc 20
||||| ||||| |||||
Db 35 CCAGACATCAAAAGCATCAC 16

```

DEFINITION Sequence 64 from patent US 5719032.
ACCESSION I88720
VERSION I88720.1 GI:3408660
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 42)
AUTHORS Vielkind,J.R.
TITLE Melanoma and prostate cancer specific antibodies for
immunodetection and immunotherapy
JOURNAL Patent: US 5719032-A 64 17-FEB-1998;
FEATURES Location/Qualifiers
source 1. .42
BASE COUNT 5 a 8 c 15 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 42;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 cctgacatcagtagcatctc 20
||| ||||| ||||| |
Db 35 CCAGACATCAAAAGCATCAC 16

RESULT 10
HUMIGHABH/c
LOCUS HUMIGHABH 45 bp mRNA PRI 09-NOV-1994
DEFINITION Human Ig rearranged mu-chain mRNA V-region, partial cds (from clone
FBM-8).
ACCESSION L06994
VERSION L06994.1 GI:184940
KEYWORDS V-region; antigen binding region; complementarity determining
region; immunoglobulin heavy chain; mu-immunoglobulin; processed
gene.
SOURCE Homo sapiens 13 week fetus bone marrow cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Ichihara,Y., Matsuoka,H. and Kurosawa,Y.
Organization of human immunoglobulin heavy chain diversity gene
loci
JOURNAL EMBO J. 7 (13), 4141-4150 (1988)
MEDLINE 89210798
AUTHORS 2 (bases 1 to 45)
Raaphorst,F.M., Sanz,I., Vossen,J.M., Schuurman,R.K.B. and van
Tol,M.J.D.
TITLE Immunoglobulin heavy chain CDR3 regions in human fetal organs:
Evidence for selection of rearrangements on the basis of CDR3
length
JOURNAL Unpublished (1993)
FEATURES Location/Qualifiers
source 1. .45
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="B lymphocyte"
/dev_stage="13 week fetus"
/tissue_type="bone marrow"
/map="14q32.33"
1. .45
/gene="IGHM"
<1. .>45
/gene="IGHM"
/note="G00-120-086"
BASE COUNT 7 a 5 c 18 g 15 t
ORIGIN

gene
V_region

Query Match 68.0%; Score 13.6; DB 97; Length 45;
Best Local Similarity 80.0%; Pred. No. 1.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 cctgacatcagtagcatctc 20
||| ||||| ||||| |
Db 35 CCAGACATCAAAAGCATCAC 16

RESULT 11
AR063732
LOCUS AR063732 30 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 17 from patent US 5846720.
ACCESSION AR063732
VERSION AR063732.1 GI:5993040
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Foulkes,J.Gordon, Liechtfried,F.E., Pieler,C., Stephenson,J.R. and
Case,C.C.
TITLE Methods of determining chemicals that modulate expression of genes
associated with cardiovascular disease
JOURNAL Patent: US 5846720-A 17 08-DEC-1998;
FEATURES Location/Qualifiers
source 1. .30
/organism="unknown"
BASE COUNT 11 a 7 c 5 g 7 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 30;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ctgacatcagtagcatct 19
||||| ||||| |||||
Db 6 CTGAATCAGCAGCACCT 23

RESULT 12
I30095
LOCUS I30095 30 bp DNA PAT 06-FEB-1997
DEFINITION Sequence 17 from patent US 5580722.
ACCESSION I30095
VERSION I30095.1 GI:1820886
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Foulkes,J.Gordon, Liechtfried,F.E., Pieler,C., Stephenson,J.R. and
Case,C.C.
TITLE Methods of determining chemicals that modulate transcriptionally
expression of genes associated with cardiovascular disease
JOURNAL Patent: US 5580722-A 17 03-DEC-1996;
FEATURES Location/Qualifiers
source 1. .30
/organism="unknown"
BASE COUNT 11 a 7 c 5 g 7 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 30;
Best Local Similarity 83.3%; Pred. No. 2e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ctgacatcagtagcatct 19
||||| ||||| |||||
Db 6 CTGAATCAGCAGCACCT 23

RESULT 13

```

```
AR020665/c
LOCUS       AR020665      23 bp      DNA
DEFINITION  Sequence 59 from patent US 5789184.
ACCESSION   AR020665
VERSION     AR020665.1  GI:3975280
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 23)
AUTHORS     Fowles,D.M., Broach,J., Manfredi,J., Klein,C., Murphy,A.J.,
            Paul,J. and Trueheart,J.
TITLE       Yeast cells engineered to produce pheromone system protein
            surrogates, and uses therefor
JOURNAL     Patent: US 5789184-A 59 04-AUG-1998;
FEATURES    Location/Qualifiers
             1..23
             /organism="unknown"
BASE COUNT  4 a      8 c      5 g      6 t
ORIGIN
Query Match      65.0%; Score 13; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 2.6e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  7 atcagtagcatct 19
Db  18 ATCAGTAGCATCT 6

RESULT 14
A51533
LOCUS       A51533      21 bp      DNA
DEFINITION  Sequence 18 from Patent EP0728842.
ACCESSION   A51533
VERSION     A51533.1  GI:2304354
KEYWORDS    unidentified.
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 21)
AUTHORS     Audonnet,J.F., Bublot,M.J., Dartell,R.J., Duinat,C.V., Laplace,E.L.
            and Riviere,M.A.
TITLE       Live recombinant avian vaccine based on an avianherpes virus,
            against Gumboro disease
JOURNAL     Patent: EP 0728842-A 18 28-AUG-1996;
            RHONE MERIEUX (FR)
COMMENT     Other publication FR 2728794 960705
            Other publication CA 2166371 960701
            Other publication AU 4063095 960711.
FEATURES    Location/Qualifiers
             1..21
             /organism="unidentified"
             /db_xref="taxon:32644"
BASE COUNT      5 a      7 c      4 g      5 t
ORIGIN

Query Match      64.0%; Score 12.8; DB 9; Length 21;
Best Local Similarity 87.5%; Pred. No. 3.3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  3 tgacatcagtagcatc 18
Db  2 TGACATCACTAGCGTC 17

RESULT 15
A51561
LOCUS       A51561      21 bp      DNA
DEFINITION  Sequence 21 from Patent EP0719864.
ACCESSION   A51561
```

```
VERSION     A51561.1  GI:2304389
KEYWORDS    unidentified.
SOURCE      unidentified.
ORGANISM    unclassified.
REFERENCE   1 (bases 1 to 21)
AUTHORS     Audonnet,J.F., Bublot,M.J., Dartell,R.J., Duinat,C.V., Laplace,E.L.
            and Riviere,M.A.
TITLE       Recombinant live avian vaccin, using an avian herpes virus as
            vector
JOURNAL     Patent: EP 0719864-A 21 03-JUL-1996;
            RHONE MERIEUX (FR)
COMMENT     Other publication FR 2728795 960705
            Other publication CA 2166367 960701
            Other publication AU 4071595 960711.
FEATURES    Location/Qualifiers
             1..21
             /organism="unidentified"
             /db_xref="taxon:32644"
BASE COUNT      5 a      7 c      4 g      5 t
ORIGIN

Query Match      64.0%; Score 12.8; DB 9; Length 21;
Best Local Similarity 87.5%; Pred. No. 3.3e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  3 tgacatcagtagcatc 18
Db  2 TGACATCACTAGCGTC 17

Search completed: October 2, 2001, 15:56:47
Job time: 14171 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:00:58 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-12
Perfect score: 20
Sequence: 1 cctgacatcagtagcatctc 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

em_esthum10:*
44: em_esthum11:*
45: em_esthum12:*
46: em_esthum13:*
47: em_esthum14:*
48: em_esthum15:*
49: em_esthum16:*
50: em_esthum17:*
51: em_esthum18:*
52: em_esthum19:*
53: em_esthum20:*
54: em_esthum21:*
55: em_esthum22:*
56: em_esthum23:*
57: em_esthum24:*
58: em_esthum25:*
59: em_esthum26:*
60: em_esthum27:*
61: em_esthum28:*
62: em_estin1:*
63: em_estin2:*
64: em_estin3:*
65: em_estin4:*
66: em_estin5:*
67: em_estom1:*
68: em_estom2:*
69: em_estov1:*
70: em_estov2:*
71: em_estpl1:*
72: em_estpl2:*
73: em_estpl3:*
74: em_estpl4:*
75: em_estpl5:*
76: em_estpl6:*
77: em_estpl7:*
78: em_estpl8:*
79: em_estpl9:*
80: em_estpl10:*
81: em_estro1:*
82: em_estro2:*
83: em_estro3:*
84: em_estro4:*
85: em_estro5:*
86: em_estro6:*
87: em_estro7:*
88: em_estro8:*
89: em_estro9:*
90: em_estro10:*
91: em_estro11:*
92: em_estro12:*
93: em_estro13:*
94: em_estro14:*
95: em_estro15:*
96: em_estro16:*
97: em_estro17:*
98: em_estro18:*
99: em_estro19:*
100: em_estro20:*
101: gb_est25:*
102: gb_est26:*
103: gb_est27:*
104: gb_est28:*
105: gb_est29:*
106: gb_est30:*
107: gb_est31:*
108: gb_est32:*
109: gb_est41:*
110: gb_est42:*
111: gb_est43:*
112: gb_est44:*
113: gb_est45:*
114: gb_est46:*
115: gb_est47:*
116:

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0149 row: A column: 07
 Seq primer: CACACGAGAAACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 19.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.2	66.0	19	250	AZ848554 2M0149A07
2	12.6	63.0	49	16	AI098615 ue31h01.y
3	12.4	62.0	31	13	AA910858 ok85c01.s
4	12	60.0	40	104	AI962543 wc53g05.x
5	12	60.0	49	14	AA975152 om39c07.s
6	11.8	59.0	34	23	AI657872 fcl4f02.y
7	11.8	59.0	41	258	HSMC04C05
8	11.8	59.0	46	249	AZ758330 IM0550115
9	11.8	59.0	49	24	AI767413 w195g10.x
10	11.6	58.0	49	2	AA087870 mn97a11.r
11	11.6	58.0	49	104	AI949145 wq08b11.x
12	11.4	57.0	47	158	H55380 CHR220319.C
13	11.2	56.0	27	249	AZ769505 IM0570L17
14	11.2	56.0	31	17	AI174159 vz84e05.r
15	11.2	56.0	31	246	AZ620204 IM0452P14
16	11.2	56.0	31	249	AZ777633 2M0012K21
17	11.2	56.0	42	244	AZ456886 IM0260C01
18	11.2	56.0	45	244	AZ465721 IM0275014
19	11.2	56.0	49	6	AA388129
20	11	55.0	34	17	AI168400 qa25e03.s
21	11	55.0	35	153	BG393308 602411521
22	11	55.0	37	14	AA995811 os05e12.s
23	11	55.0	37	17	AI208016 qg60f11.x
24	11	55.0	43	190	W62348 md99c05.r1
25	11	55.0	50	107	AU103398 AU103398
26	11	55.0	50	107	AU104027 AU104027
27	11	55.0	50	107	AU107009 AU107009
28	10.8	54.0	22	244	AZ459715 IM0264E13
29	10.8	54.0	26	250	AZ814137 2M0081B16
30	10.8	54.0	34	250	AZ856611 2M0161P06
31	10.8	54.0	39	104	AJ239886 AJ239886
32	10.8	54.0	43	246	AZ592784 IM0403024
33	10.8	54.0	43	258	TA5F10Q
34	10.8	54.0	46	189	T53599 ya97f09.s1
35	10.6	53.0	28	17	AI187132 qe42f11.x
36	10.6	53.0	30	242	AZ345667 IM0080022
37	10.6	53.0	34	17	AI188358 qd08c03.x
38	10.6	53.0	35	242	AZ389531 IM0150P07
39	10.6	53.0	36	247	AZ659063 IM0536N10
40	10.6	53.0	39	191	Z20226 HSAABOLR.S
41	10.6	53.0	40	14	AA993748 os54g04.s
42	10.6	53.0	40	106	AU013065 AU013065
43	10.6	53.0	40	106	AU013075 AU013075
44	10.6	53.0	40	139	BE739544 601556485
45	10.6	53.0	40	244	AZ462628 IM0269E08

ALIGNMENTS

RESULT 1
 AZ848554/c
 LOCUS 2M0149A07R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC2M0149A07 R, DNA sequence.
 AZ848554
 VERSION GSS
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 19)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.

BASE COUNT 2 a 4 c 7 g 6 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 250; Length 19;
 Best Local Similarity 83.3%; Pred. No. 1.7e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 cctgcacatcagtagcatc 18
 ||||| |||||
 Db 19 CCTGACGACCTAGCAGC 2

RESULT 2
 AI098615

LOCUS ue31h01.y1 Sugano mouse liver mlia Mus musculus cDNA clone
 DEFINITION IMAGE:1482001 5', similar to TR:P97298 SPROMAL CELL DERIVED
 FACTOR 3 ;, mRNA sequence.

ACCESSION AI098615
 VERSION AI098615.1 GI:3448140
 KEYWORDS EST.
 SOURCE house mouse.

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 49)
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 The RASHU-HHMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 Washington University EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:930357

Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: custom primer used
 High quality sequence stop: 1.

Location/Qualifiers
 1. .49
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone_image="IMAGE:1482001"
 /clone_lib="Sugano mouse liver mlia"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII (CAGCTGTG); Site_2: DraIII (CAGCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGCGCTTTTGTGTGTGTGT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTTGGCCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CAGCTGTG, 3' site CAGCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCCTGCTCTCTAAAGCTGGC and 3' end primer CGACCTGCAGCTCGACACA."

FEATURES

source

BASE COUNT 13 a 18 c 10 g 8 t
 ORIGIN
 Query Match 63.0%; Score 12.6; DB 16; Length 49;
 Best Local Similarity 78.9%; Pred. No. 3.8e+04;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 cctgacatcagtagcatct 19
 ||||| |||||
 Db 23 CCTGATCCACAGCACCT 41
 RESULT 3
 AA910858 31 bp mRNA EST 13-APR-1998
 LOCUS OK85c01.s1 NCI_CGAP_kid3 Homo sapiens cDNA clone IMAGE:1520736 3'
 DEFINITION similar to TR:Q13531 Q13531 DISHEVELLED ;, mRNA sequence.
 ACCESSION AA910858
 VERSION AA910858.1 GI:3050148
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 31)
 REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Seq primer: -40ml3 fwd. ET from Amersham
 High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers
 1. .31
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:1520736"
 /clone_lib="NCI_CGAP_kid3"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. mRNA source: 2 pooled kidneys. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 4 a 6 c 16 g 5 t
 ORIGIN

Query Match 62.0%; Score 12.4; DB 13; Length 31;
 Best Local Similarity 92.9%; Pred. No. 4.5e+04;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 gacatcagtagcat 17
 ||||| |||||

Db 9 GGCATCAGTAGCAT 22

RESULT 4

AI962543

LOCUS

DEFINITION

AI962543 40 bp mRNA EST 08-MAR-2000

similar to TR:O88247 O88247 MSZF52 ;, mRNA sequence.

ACCESSION AI962543

VERSION AI962543.1 GI:5755256

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 40)

REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

TITLE Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Insert Length: 884 Std Error: 0.00

double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

BASE COUNT 13 a 7 c 11 g 3 t
 ORIGIN
 Query Match 59.0%; Score 11.8; DB 23; Length 34;
 Best Local Similarity 86.7%; Pred. No. 9.1e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gacatcagtagcgc 18

||||||| |||||

Db 2 GACATCAGCAGCAGC 16

RESULT 7

HSMC04C05 41 bp DNA GSS 29-MAY-1997
 LOCUS H.sapiens DNA for trapped exon (ID RMC04C05), genomic survey
 DEFINITION sequence.
 X88541
 X88541.1 GI:1437552
 GSS.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 41)
 AUTHORS Chen,H.M., Rossier,C., Chrast,R. and Antonarakis,S.E.
 TITLE Cloning of trapped exons from human chromosome 21
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 41)
 AUTHORS Antonarakis,S.E.
 TITLE Direct Submission
 JOURNAL Submitted (17-MAR-1995) Stylianos E. Antonarakis, Division of Medical Genetics, University and Cantonal Hospital of Geneva, CMU, 1 rue Michel-Servet, 1211 Geneva, SWITZERLAND

REFERENCE 3 (bases 1 to 41)
 AUTHORS Chen,H., Chrast,R., Rossier,C., Morris,M.A., Lalioti,M.D. and Antonarakis,S.E.
 TITLE Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping
 JOURNAL Genome Res. 6 (8), 747-760 (1996)
 MEDLINE 97011340

FEATURES Location/Qualifiers
 source 1..41
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="21"
 exon 1..41
 /note="trapped exon"

BASE COUNT 12 a 8 c 6 g 12 t 3 others
 ORIGIN
 Query Match 59.0%; Score 11.8; DB 258; Length 41;
 Best Local Similarity 86.7%; Pred. No. 9.4e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 catcagtagcgc 20

Db 14 CAACAGGAGCATCTC 28
 || ||| ||||| |||||

RESULT 8

AZ758330/c 46 bp DNA GSS 16-FEB-2001
 LOCUS 1M050115F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC1M050115 F, DNA sequence.

ACCESSION AZ758330
 VERSION AZ758330.1 GI:12864016

KEYWORDS GSS.
 SOURCE house mouse.

ORGANISM

REFERENCE 1 (bases 1 to 46)
 AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunne@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0550 row: 1 column: 15
 Seq primer: CCTGTAAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 46.

FEATURES

Location/Qualifiers

source 1..46

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M050115"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 12 a 6 c 11 g 17 t

ORIGIN

Query Match 59.0%; Score 11.8; DB 249; Length 46;

Best Local Similarity 86.7%; Pred. No. 9.5e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 gacatcagtagcatc 18
 |||||
 Db 18 GACATCAGTAGAACC 4

RESULT 9
 A1767413/c
 LOCUS
 DEFINITION A1767413 49 bp mRNA EST 21-DEC-1999
 w195g10.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2401122 3'
 similar to WP:F55F8.5 Cell1196 ;, mRNA sequence.
 A1767413
 ACCESSION A1767413.1 GI:5233842
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 49)
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert Length: 969 Std Error: 0.00
 Seq primer: 40UP from Gibco
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 49
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2401122"
 /clone_lib="NCI_CGAP_Kid12"
 /tissue_type="2 pooled tumors (clear cell type)"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Plasmid DNA from the normalized library NCI_CGAP_Kid5 was
 prepared, and ss circles were made in vitro. Following HAP
 purification, this DNA was used as tracer in a subtractive
 hybridization reaction. The driver was PCR-amplified cDNAs
 from a pool of 5,000 clones made from the same library
 (cloneIDs 1323912-1325831, 1471368-1472903 and
 1492104-1493255). Subtraction by Bento Soares and M.
 Fatima Bonaldo."

BASE COUNT 12 a 11 c 12 g 14 t
 ORIGIN

Query Match 59.0%; Score 11.8; DB 24; Length 49;
 Best Local Similarity 86.7%; Pred. No. 9.6e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 tgacatcagtagcatc 17
 |||||
 Db 36 TGACATCAGTAAAT 22

RESULT 10
 AA087870
 LOCUS
 DEFINITION mn97all.r1 Stratagene mouse lung 937302 EST 15-FEB-1997
 IMAGE:551996 5' similar to TR:G435967 G435967 APOLIPOPROTEIN F. ;,

mrna sequence.
 ACCESSION AA087870
 VERSION AA087870.1 GI:1631078
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 house mouse.
 REFERENCE 1 (bases 1 to 49)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Scheilenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:332788
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 49
 /organism="Mus musculus"
 /strain="C57BL/6 x CBA"
 /db_xref="taxon:10090"
 /clone="IMAGE:551996"
 /clone_lib="Stratagene mouse lung 937302"
 /sex="female"
 /tissue_type="lung"
 /dev_stage="6-8 month old"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
 ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
 dt. 6-8 month old female lung and 1.5 year old male lung
 were source of mrna. Average insert size: 1.5 kb; Uni-ZAP
 XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3'
 adaptor sequence: 5' CTCAGTTTTTTTTTTTTTT 3'"

BASE COUNT 10 a 17 c 8 g 14 t
 ORIGIN

Query Match 58.0%; Score 11.6; DB 2; Length 49;
 Best Local Similarity 77.8%; Pred. No. 1.2e+05;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 cctgacatcagtagcatc 18
 |||||
 Db 2 CCTGACATGCATAGGCTC 19

RESULT 11
 A1949145
 LOCUS
 DEFINITION A1949145 49 bp mRNA EST 08-MAR-2000
 wq08b11.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2470653 3'
 similar to TR:O60455 O60455 POLYADENYLATE BINDING
 PROTEIN-INTERACTING PROTEIN-1. ;, mRNA sequence.
 A1949145
 ACCESSION A1949145.1 GI:5741455
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 49)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert Length: 779 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.

FEATURES

source
 1. .49
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2470653"
 /clone_lib="NCI_CGAP_Kid12"
 /tissue_type="2 pooled tumors (clear cell type)"
 /lab_host="DH10B"
 /note="Organ: kidney; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kid5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaudo."

BASE COUNT 10 a 13 c 9 g 16 t 1 others
 ORIGIN

Query Match 58.0%; Score 11.6; DB 104; Length 49;
 Best Local Similarity 77.8%; Pred. No. 1.2e+05;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ctgacatcagtagcatct 19
 |||||
 Db 18 CTTACTGCAGTCGCACT 35

RESULT 12

H53380 47 bp mRNA EST 07-DEC-1995
 LOCUS CHR220319 Chromosome 22 exon Homo sapiens cDNA clone C22_405 5', mRNA sequence.

ACCESSION H53380
 VERSION H53380.1 GI:1108246
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 47)
 Trofatter, J.A., Long, K.R., Murrell, J.R., Stotler, C.J., Gusella, J.F. and Buckler, A.J.

TITLE An expression-independent catalog of genes from human chromosome 22
JOURNAL Genome Res. 5 (3), 214-224 (1995)
MEDLINE 96159527

COMMENT Contact: Buckler AJ
 Molecular Neurogenetics Unit
 Massachusetts General Hospital
 Building 149, 13th St., Charlestown MA 02129
 Tel: 617249616

FEATURES

source

Location/Qualifiers
 1. .47
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="C22_405"
 /clone_lib="Chromosome 22 exon"
 /lab_host="E. coli DH5a"
 /note="Vector: pBluescriptIIKS+; Site_1: Sal I; Site_2: Bam HI (destroyed); Exons were isolated from human chromosome 22 specific cosmids using a modification of the method of exon amplification (Proc. Natl. Acad. Sci. USA 88:4005-4009, 1991). Amplified exons were digested with Sal I and Bgl II and subsequently cloned into pBluescriptIIKS+ at the Sal I and Bam HI sites."

BASE COUNT 15 a 11 c 12 g 9 t
 ORIGIN

Query Match 57.0%; Score 11.4; DB 158; Length 47;
 Best Local Similarity 92.3%; Pred. No. 1.5e+05;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 atcagtagcatct 19
 |||||
 Db 31 ATCATTAGCATCT 43

RESULT 13

AZ769505 27 bp DNA GSS 16-FEB-2001
 LOCUS 1M0570117F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0570117 F, DNA sequence.

ACCESSION AZ769505
 VERSION AZ769505.1 GI:12889709
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 27)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL

COMMENT Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0570 row: L column: 17

Seq primer: CCGTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 27.

FEATURES

source

Location/Qualifiers
 1. .27
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0570117"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, P-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g114732114 [gb|AF129072.1]), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```

BASE COUNT      9 a      3 c      8 g      7 t
ORIGIN

Query Match      56.0%; Score 11.2; DB 249; Length 27;
Best Local Similarity 81.2%; Pred. No. 1.8e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 2 ctgacatcagtagcat 17
    ||| ||||| ||| ||
Db 5 CTGTTATCAGTAGGAT 20

```

```

RESULT 14
A1174159
LOCUS      A1174159      31 bp      mRNA      EST      07-OCT-1998
DEFINITION v284e05.rl Soares_mammary_gland_NbMMG Mus musculus cDNA clone
IMAGE:1333184 5' similar to SW:NUIM_MOUSE P03888 NADH-UBIQUINONE
OXIDOREDUCTASE CHAIN 1 ; mRNA sequence.

```

```

ACCESSION      A1174159.1 GI:3720301
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 31)
Maria.M., Hillier.L., Allen.M., Bowles.M., Dietrich.N., Dubuque.T.,
Geisel.S., Kucaba.T., Lacy.M., Le.M., Martin.J., Morris.M.,
Schellenberg.K., Steptoe.M., Tan.F., Underwood.K., Moore.B.,
Theising.B., Wylie.T., Lennon.G., Soares.B., Wilson.R. and
Waterston.R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:692728

```

```

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .31
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1333184"
/clone_lib="Soares_mammary_gland_NbMMG"

```

```

FEATURES
source

```

```

/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
) with a modified polylinker; Site_1: Not I; Site_2: Eco
RI; 1st strand cDNA was primed with a Not I - oligo(dT)
primer [5'
TGTATCAACATCTGAAGTGGAGCGCGCGAATGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

```

```

BASE COUNT      9 a      8 c      7 g      7 t
ORIGIN

Query Match      56.0%; Score 11.2; DB 17; Length 31;
Best Local Similarity 81.2%; Pred. No. 1.8e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 1 cctgacatcagtagca 16
    ||| |||| |||| |
Db 13 CCTAACATTAGTAGAA 28

```

```

RESULT 15
A2620204
LOCUS      A2620204      31 bp      DNA      GSS      13-DEC-2000
DEFINITION 1M0452P14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0452P14 R, DNA sequence.

```

```

ACCESSION      A2620204
VERSION      A2620204.1 GI:11742394
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 31)
Dunn.D., Aoyagi.A., Barber.M., Beacorn.T., Duval.B., Hamill.C.,
Islam.H., Longacre.S., Mahmoud.M., Meenen.E., Pedersen.T., Reilly
,M., Rose.M., Rose.R., Stokes.R., Tingey.A., von Niederhausern.A.
and Wright.D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center

```

```

JOURNAL
COMMENT
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0452 row: P column: 14
Seq primer: CACAGGAAACAGCATGACC
Class: plasmid ends
High quality sequence stop: 31.
Location/Qualifiers
1. .31
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0452P14"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource

```

```

FEATURES
source

```

(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 4 a 5 c 16 g 6 t
ORIGIN

Query Match 56.0%; Score 11.2; DB 246; Length 31;
Best Local Similarity 81.2%; Pred. No. 1.8e+05;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ctgacatcagtagcat 17
||||| |||||
Db 3 CTGACAGCAGTGGCTT 18

Search completed: October 2, 2001, 15:01:00
Job time: 10829 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:47 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-14

Perfect score: 20

Sequence: 1 ggttaggatgtgcgcgtca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba1.*

2: gb_ba2.*

3: gb_ba3.*

4: gb_in1.*

5: gb_in2.*

6: gb_in3.*

7: gb_om.*

8: gb_ov.*

9: gb_pat1.*

10: gb_pat2.*

11: gb_ph.*

12: gb_pl1.*

13: gb_pl2.*

14: gb_pl3.*

15: gb_pl4.*

16: em_ba1.*

17: em_ba2.*

18: em_fun.*

19: em_htgo_hum.*

20: em_htgo_inv.*

21: em_htgo_rod.*

22: em_htg_hum1.*

23: em_htg_hum2.*

24: em_htg_hum3.*

25: em_htg_hum4.*

26: em_htg_hum5.*

27: em_htg_hum6.*

28: em_htg_hum7.*

29: em_htg_hum8.*

30: em_htg_inv1.*

31: em_htg_inv2.*

32: em_htg_other.*

33: em_htg_rod.*

34: em_hum1.*

35: em_hum2.*

36: em_hum3.*

37: em_hum4.*

38: em_hum5.*

39: em_hum6.*

40: em_hum7.*

41: em_in.*

42: em_om.*

43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vil2.*
60: gb_vil2.*
61: gb_vil2.*
62: gb_vil2.*
63: gb_vil2.*
64: gb_vil2.*
65: gb_vil2.*
66: gb_vil2.*
67: gb_vil2.*
68: gb_vil2.*
69: gb_vil2.*
70: gb_vil2.*
71: gb_vil2.*
72: gb_vil2.*
73: gb_vil2.*
74: gb_vil2.*
75: gb_vil2.*
76: gb_vil2.*
77: gb_vil2.*
78: gb_vil2.*
79: gb_vil2.*
80: gb_vil2.*
81: gb_vil2.*
82: gb_vil2.*
83: gb_vil2.*
84: gb_vil2.*
85: gb_vil2.*
86: gb_vil2.*
87: gb_vil2.*
88: gb_vil2.*
89: gb_vil2.*
90: gb_vil2.*
91: gb_vil2.*
92: gb_vil2.*
93: gb_vil2.*
94: gb_vil2.*
95: gb_vil2.*
96: gb_vil2.*
97: gb_vil2.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	13.8	69.0	24	9 AR016339	AR016339 Sequence 18
C 2	13.8	69.0	24	10 I59963	I59963 Sequence 18
C 3	13.8	69.0	24	10 I86044	I86044 Sequence 18
C 4	13.8	69.0	30	10 I24249	I24249 Sequence 36
C 5	12.2	61.0	34	9 AX001078	AX001078 Sequence
6	12	60.0	26	9 AR087980	AR087980 Sequence 87
7	12	60.0	26	10 I80198	I80198 Sequence 87
8	12	60.0	32	10 I07710	I07710 Sequence 15

9 12 60.0 39 9 A92270
c 10 12 60.0 42 95 S77075
c 11 12 60.0 45 97 S63845
c 12 11.8 59.0 20 10 E08846
c 13 11.8 59.0 26 9 A087898
c 14 11.8 59.0 26 9 A087920
c 15 11.8 59.0 26 9 A087948
c 16 11.8 59.0 26 9 A087958
c 17 11.8 59.0 26 9 A087960
c 18 11.8 59.0 26 9 A087974
c 19 11.8 59.0 26 9 A087976
c 20 11.8 59.0 26 9 A087978
c 21 11.8 59.0 26 9 A090813
c 22 11.8 59.0 26 10 I80116
c 23 11.8 59.0 26 10 I80138
c 24 11.8 59.0 26 10 I80166
c 25 11.8 59.0 26 10 I80176
c 26 11.8 59.0 26 10 I80178
c 27 11.8 59.0 26 10 I80192
c 28 11.8 59.0 26 10 I80194
c 29 11.8 59.0 26 10 I80196
c 30 11.8 59.0 30 10 I24263
c 31 11.8 59.0 35 9 A69153
c 32 11.8 59.0 40 97 H5274584
c 33 11.8 59.0 46 9 AX059019
c 34 11.8 59.0 50 10 E16496
c 35 11.6 58.0 23 9 AX003306
c 36 11.6 58.0 24 9 A083193
c 37 11.6 58.0 25 9 A090589
c 38 11.6 58.0 26 9 A090376
c 39 11.6 58.0 36 9 A32693
c 40 11.6 58.0 36 9 AX036277
c 41 11.6 58.0 36 9 AX038887
c 42 11.6 58.0 48 9 A053527
c 43 11.4 57.0 18 9 A0909347
c 44 11.4 57.0 18 9 A060812
c 45 11.4 57.0 18 10 I31591

ALIGNMENTS

RESULT 1
AR016339/c
LOCUS AR016339 24 bp DNA
DEFINITION Sequence 18 from patent US 5776694.
ACCESSION AR016339
VERSION AR016339.1 GI:3972616
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 24)
AUTHORS Sheiness,D.K., Adams,T.H., Stamm,M.R., Cangelosi,G.A.,
Britschgi,T.B. and Dix,C.K.
TITLE Diagnostic kits useful for selectively detecting microorganisms in
samples
JOURNAL Patent: US 5776694-A 18 07-JUL-1998;
FEATURES Location/Qualifiers
source 1..24
BASE COUNT 7 a 7 c 6 g 4 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 24;
Best Local Similarity 88.2%; Pred. No. 1.8e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 ttaggatggtgccgc 19
||| ||||| |||||
Db 18 TTGGGATGGAGCCGTC 2

RESULT 2
I59963/c
LOCUS I59963 24 bp DNA
DEFINITION Sequence 18 from patent US 5654418.
ACCESSION I59963
VERSION I59963.1 GI:2478595
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 24)
AUTHORS Sheiness,D.K., Cangelosi,G.A. and Britschgi,T.B.
TITLE Nucleic acid probes useful for detecting microorganisms associated
with vaginal infections
JOURNAL Patent: US 5654418-A 18 05-AUG-1997;
FEATURES Location/Qualifiers
source 1..24
BASE COUNT 7 a 7 c 6 g 4 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 24;
Best Local Similarity 88.2%; Pred. No. 1.8e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ttaggatggtgccgc 19
||| ||||| |||||
Db 18 TTGGGATGGAGCCGTC 2

RESULT 3
I86044/c
LOCUS I86044 24 bp DNA
DEFINITION Sequence 18 from patent US 5700636.
ACCESSION I86044
VERSION I86044.1 GI:3205762
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 24)
AUTHORS Sheiness,D.K., Adams,T.H., Stamm,M.R., Cangelosi,G.A.,
Britschgi,T.B. and Dix,C.K.
TITLE Methods for selectively detecting microorganisms associated with
vaginal infections in complex biological samples
JOURNAL Patent: US 5700636-A 18 23-DEC-1997;
FEATURES Location/Qualifiers
source 1..24
BASE COUNT 7 a 7 c 6 g 4 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 24;
Best Local Similarity 88.2%; Pred. No. 1.8e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ttaggatggtgccgc 19
||| ||||| |||||
Db 18 TTGGGATGGAGCCGTC 2

RESULT 4
I24249
LOCUS I24249 30 bp DNA
DEFINITION Sequence 36 from patent US 5543293.
ACCESSION I24249
VERSION I24249.1 GI:1604119
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

```
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Gold,L. and Tasset,D.
TITLE DNA ligands of thrombin
JOURNAL Patent: US 5343293-A 36 06-AUG-1996;
FEATURES Location/Qualifiers
source 1..30
BASE COUNT 4 a 1 c 15 g 10 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 10; Length 30;
Best Local Similarity 88.2%; Pred. No. 1.7e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ggtaggagtggtgcgcg 17
||| ||||| ||||| |||
Db 11 GGTTAGGATGGTAGCG 27

RESULT 5
AX001078/c 34 bp DNA PAT 10-MAR-2000
LOCUS Sequence 3 from Patent WO9902647.-
ACCESSION AX001078
VERSION AX001078.1 GI:7241296
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 34)
AUTHORS Sandig,V. and Loeser,P.
TITLE CLONING VECTORS FOR PRODUCING ADENOVIRAL MINIMAL VIRUSES
JOURNAL Patent: WO 9902647-A 3 21-JAN-1999;
FEATURES HEPAVEC AG FUER GENTHERAPIE (DE); SANDIG VOLKER (DE)
source 1..34
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 8 a 11 c 8 g 7 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 34;
Best Local Similarity 82.4%; Pred. No. 1.2e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 taggagtggtgcgcgta 20
||| ||||| |||||
Db 18 TAGAGTTGGTACCGTCA 2

RESULT 6
AR087980 26 bp DNA PAT 07-SEP-2000
LOCUS Sequence 87 from patent US 5989821.
ACCESSION AR087980
VERSION AR087980.1 GI:10014743
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 26)
AUTHORS Goh,S.Han, Chow,A. and Hemmingsen,S.
TITLE Universal targets for species identification
JOURNAL Patent: US 5989821-A 87 23-NOV-1999;
FEATURES Location/Qualifiers
source 1..26
/organism="unknown"
BASE COUNT 5 a 6 c 10 g 5 t
ORIGIN

Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Gold,L. and Tasset,D.
TITLE DNA ligands of thrombin
JOURNAL Patent: US 5343293-A 36 06-AUG-1996;
FEATURES Location/Qualifiers
source 1..30
BASE COUNT 4 a 1 c 15 g 10 t
ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 26;
Best Local Similarity 75.0%; Pred. No. 1.6e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggtaggagtggtgcgcgta 20
||| ||||| ||||| |||
Db 6 GGCTGGGATGGTACCACCA 25

RESULT 7
I80198 26 bp DNA PAT 10-JUN-1998
LOCUS Sequence 87 from patent US 5708160.
ACCESSION I80198
VERSION I80198.1 GI:3208488
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 26)
AUTHORS Goh,S.Han, Chow,A.W. and Hemmingsen,S.
TITLE HSP-60 genomic locus and primers for species identification
JOURNAL Patent: US 5708160-A 87 13-JAN-1998;
FEATURES Location/Qualifiers
source 1..26
/organism="unknown"
BASE COUNT 5 a 6 c 10 g 5 t
ORIGIN

Query Match 60.0%; Score 12; DB 10; Length 26;
Best Local Similarity 75.0%; Pred. No. 1.6e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggtaggagtggtgcgcgta 20
||| ||||| ||||| |||
Db 6 GGCTGGGATGGTACCACCA 25

RESULT 8
I07710 32 bp PAT 02-DEC-1994
LOCUS Sequence 15 from Patent EP 0362526.
ACCESSION I07710
VERSION I07710.1 GI:589756
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 32)
AUTHORS Springer,T.A. and Larson,R.
TITLE The Alpha-subunit of the LFA-1 leukocyte adhesion receptor
JOURNAL Patent: EP 0362526-A2 15 11-APR-1990;
FEATURES Location/Qualifiers
source 1..32
/organism="unknown"
BASE COUNT 5 a 3 c 14 g 10 t
ORIGIN

Query Match 60.0%; Score 12; DB 10; Length 32;
Best Local Similarity 75.0%; Pred. No. 1.6e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggtaggagtggtgcgcgta 20
||| ||||| ||||| |||
Db 11 GGTCATGGATGGTGGGCTCA 30

RESULT 9
A92270 39 bp DNA PAT 23-JAN-2000
LOCUS
```

```

DEFINITION Sequence 7 from Patent WO9818943.
ACCESSION A92270
VERSION A92270.1 GI:6741040
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 39)
AUTHORS Kajava,A. and Cramer,R.
TITLE METHOD FOR THE OLIGOMERISATION OF PEPTIDES
JOURNAL Patent: WO 9818943-A 7 07-MAY-1998;
CIBA GEIGY AG (CH); KAJAVA ANDREY (CH)
FEATURES
    source
        1..39
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT 8 a 14 c 13 g 4 t
ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 39;
Best Local Similarity 75.0%; Pred. No. 1.6e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ggttaggagtggtgcgcgtca 20
    ||||| ||| ||||| |||
Db 17 GGTACGGCAAGCGCGCTCA 36

RESULT 10
S77075/c
LOCUS
DEFINITION T-cell receptor beta chain V-J region (CDR3 region) [mice,
transgenic, mRNA Partial, 42 nt].
ACCESSION S77075
VERSION S77075.1 GI:242711
KEYWORDS
SOURCE
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Jorgensen,J.L., Esser,U., de St Groth,B.F., Reay,P.A. and
Davis,M.M.
TITLE Mapping T-cell receptor-peptide contacts by variant peptide
immunization of single-chain transgenics
JOURNAL Nature 355 (6357), 224-230 (1992)
MEDLINE 92114961
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsg 77075] from the original journal article.
This sequence comes from FIG. 4 c.
FEATURES
    source
        1..42
            Location/Qualifiers
            /organism="Mus sp."
            /db_xref="taxon:10095"
            <1..>42
            /gene="T-cell receptor beta chain V-J region"
            /note="This sequence comes from FIG. 4 c: conceptual
            translation presented here differs from translation in
            publication"
            /codon_start=1
            /product="T-cell receptor beta chain V-J region"
            /protein_id="AAB20961.2"
            /db_xref="GI:9027588"
            /translation="CAGTVRDQYAEQFF"
BASE COUNT 7 a 13 c 12 g 10 t
ORIGIN

Query Match 60.0%; Score 12; DB 95; Length 42;
Best Local Similarity 75.0%; Pred. No. 1.5e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

DEFINITION Sequence 7 from Patent WO9818943.
ACCESSION A92270
VERSION A92270.1 GI:6741040
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 39)
AUTHORS Kajava,A. and Cramer,R.
TITLE METHOD FOR THE OLIGOMERISATION OF PEPTIDES
JOURNAL Patent: WO 9818943-A 7 07-MAY-1998;
CIBA GEIGY AG (CH); KAJAVA ANDREY (CH)
FEATURES
    source
        1..39
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT 8 a 14 c 13 g 4 t
ORIGIN

Query Match 60.0%; Score 12; DB 9; Length 39;
Best Local Similarity 75.0%; Pred. No. 1.6e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ggttaggagtggtgcgcgtca 20
    ||||| ||| ||||| |||
Db 17 GGTACGGCAAGCGCGCTCA 36

RESULT 10
S77075/c
LOCUS
DEFINITION T-cell receptor beta chain V-J region (CDR3 region) [mice,
transgenic, mRNA Partial, 42 nt].
ACCESSION S77075
VERSION S77075.1 GI:242711
KEYWORDS
SOURCE
ORGANISM Mus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Jorgensen,J.L., Esser,U., de St Groth,B.F., Reay,P.A. and
Davis,M.M.
TITLE Mapping T-cell receptor-peptide contacts by variant peptide
immunization of single-chain transgenics
JOURNAL Nature 355 (6357), 224-230 (1992)
MEDLINE 92114961
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsg 77075] from the original journal article.
This sequence comes from FIG. 4 c.
FEATURES
    source
        1..42
            Location/Qualifiers
            /organism="Mus sp."
            /db_xref="taxon:10095"
            <1..>42
            /gene="T-cell receptor beta chain V-J region"
            /note="This sequence comes from FIG. 4 c: conceptual
            translation presented here differs from translation in
            publication"
            /codon_start=1
            /product="T-cell receptor beta chain V-J region"
            /protein_id="AAB20961.2"
            /db_xref="GI:9027588"
            /translation="CAGTVRDQYAEQFF"
BASE COUNT 7 a 13 c 12 g 10 t
ORIGIN

Query Match 60.0%; Score 12; DB 95; Length 42;
Best Local Similarity 75.0%; Pred. No. 1.5e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ggttaggagtggtgcgcgtca 20
    ||||| ||| ||||| |||
Db 22 GGTCCCGAGCGGTGCCGCA 3

RESULT 11
S63845/c
LOCUS
DEFINITION (D3(TM4-del))-D3 dopamine receptor {3' region, alternatively
spliced} [human, peripheral blood lymphocytes, mRNA Partial Mutant,
45 nt].
ACCESSION S63845
VERSION S63845.1 GI:400540
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Nagai,Y., Ueno,S., Saeki,Y., Soga,F. and Yanagihara,T.
TITLE Expression of the D3 dopamine receptor gene and a novel variant
transcript generated by alternative splicing in human peripheral
blood lymphocytes
JOURNAL Biochem. Biophys. Res. Commun. 194 (1), 368-374 (1993)
MEDLINE 93326145
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsg 135517] from the original journal article.
This sequence comes from Fig. 4.
FEATURES
    source
        1..45
            Location/Qualifiers
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /cell_type="peripheral lymphocytes"
            /tissue_type="blood"
            <1..>45
            /gene="D3R"
            /allele="D3(TM4-del)"
            <1..>45
            /gene="D3R"
            /note="This sequence comes from Fig. 4; conceptual
            translation presented here differs from translation in
            publication"
            /codon_start=1
            /product="D3 dopamine receptor"
            /protein_id="AAB27543.2"
            /db_xref="GI:7704924"
            /translation="SIDRGPCLLLHQP"
BASE COUNT 11 a 17 c 9 g 8 t
ORIGIN

Query Match 60.0%; Score 12; DB 97; Length 45;
Best Local Similarity 75.0%; Pred. No. 1.5e+05;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ggttaggagtggtgcgcgtca 20
    ||||| ||| ||||| |||
Db 41 GGTTCGAGATGGAGCAGACA 22

RESULT 12
E08846/c
LOCUS
DEFINITION PCR primer to detect polymorphism of Histamine H1 receptor gene.
ACCESSION E08846
VERSION E08846.1 GI:2176950
KEYWORDS
SOURCE JP 1995067654-A/4.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Shimizu,S., Shimazu,M., Habano,W. and Hayashi,T.
TITLE HUMAN HISTAMINE H1 RECEPTOR GENE AND ITS UTILIZATION

```

JOURNAL Patent: JP 1995067654-A 4 14-MAR-1995;

COMMENT MITSUBISHI KAGAKU B C L:KK

OS None

OC Artificial sequences.

PN JP 1995067654-A/4

PD 14-MAR-1995

PF 03-SEP-1993 JP 1993219544

PI SHIMIZU SHOICHI, SHIMAZU MITSUNOBU, HABANO WATARU, PI

HAYASHI TOMOKO

PC C12N15/09,C12Q1/68;

CC strandedness: Single;

CC topology: Linear;

FH Key Location/Qualifiers

FT source

FT source 1..20 /organism='Artificial sequences'.

FEATURES Location/Qualifiers

source 1..20

BASE COUNT 5 a 7 c 3 g 5 t

ORIGIN /db_xref='taxon:32644'

Query Match 59.0%; Score 11.8; DB 10; Length 20;

Best Local Similarity 86.7%; Pred. No. 2.2e+05;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 gggatgggtgcgcgtca 20

DB 19 GGGATGGTGACTTCA 5

RESULT 13

AR087898/c

LOCUS AR087898 26 bp DNA PAT 07-SEP-2000

DEFINITION Sequence 5 from patent US 5989821.

ACCESSION AR087898

VERSION AR087898.1 GI:10014661

KEYWORDS

SOURCE

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 26)

AUTHORS Goh,S.Han, Chow,A. and Hemmingsen,S.

TITLE Universal targets for species identification

JOURNAL Patent: US 5989821-A 5 23-NOV-1999;

FEATURES Location/Qualifiers

source 1..26

BASE COUNT 6 a 11 c 7 g 2 t

ORIGIN /organism='unknown'

Query Match 59.0%; Score 11.8; DB 9; Length 26;

Best Local Similarity 86.7%; Pred. No. 2.1e+05;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 gggatgggtgcgcgtca 20

DB 26 GTGGTGGTCCCGTCA 12

RESULT 14

AR087920/c

LOCUS AR087920 26 bp DNA PAT 07-SEP-2000

DEFINITION Sequence 27 from patent US 5989821.

ACCESSION AR087920

VERSION AR087920.1 GI:10014683

KEYWORDS

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 26)

AUTHORS Goh,S.Han, Chow,A. and Hemmingsen,S.

TITLE Universal targets for species identification

JOURNAL Patent: US 5989821-A 27 23-NOV-1999;

FEATURES Location/Qualifiers

source 1..26

BASE COUNT 6 a 11 c 7 g 2 t

ORIGIN /organism='unknown'

Query Match 59.0%; Score 11.8; DB 9; Length 26;

Best Local Similarity 86.7%; Pred. No. 2.1e+05;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 gggatgggtgcgcgtca 20

DB 26 GTGGTGGTCCCGTCA 12

RESULT 15

AR087948/c

LOCUS AR087948 26 bp DNA PAT 07-SEP-2000

DEFINITION Sequence 55 from patent US 5989821.

ACCESSION AR087948

VERSION AR087948.1 GI:10014711

KEYWORDS

SOURCE

Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 26)

AUTHORS Goh,S.Han, Chow,A. and Hemmingsen,S.

TITLE Universal targets for species identification

JOURNAL Patent: US 5989821-A 55 23-NOV-1999;

FEATURES Location/Qualifiers

source 1..26

BASE COUNT 5 a 9 c 9 g 3 t

ORIGIN /organism='unknown'

Query Match 59.0%; Score 11.8; DB 9; Length 26;

Best Local Similarity 86.7%; Pred. No. 2.1e+05;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 gggatgggtgcgcgtca 20

DB 26 GTGGTGGTCCCGTCA 12

Search completed: October 2, 2001, 15:56:48

Job time: 14172 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:41 ; Search time 876.95 Seconds
(without alignments)
14.320 Million cell updates/sec

Title: US-09-757-100B-14

Perfect score: 20

Sequence: 1 ggtaggagtggtgcgcgtca 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	22 AAC65546	Human focal adhesi
2	15	75.0	15	22 AAC65546	Human focal adhesi
3	14.2	71.0	34	19 AAV20799	Flea serine protea
C 4	13.8	69.0	24	13 AAQ24636	Probe Gv003 specif
C 5	13.8	69.0	24	14 AAQ52752	Sequence of probe
C 6	13.8	69.0	24	18 AAT89823	Oligonucleotide pr
C 7	13.8	69.0	24	19 AAV36792	Nucleotide sequenc
C 8	13.8	69.0	30	16 AAT00210	Thrombin DNA ligand
C 9	13.8	69.0	30	22 AAF0762	Thrombin high affi
C 10	13.4	67.0	18	20 AAZ40968	Human Rhoc phospho
C 11	13.4	67.0	18	20 AAZ40969	Human Rhoc phospho

C 12	13.4	67.0	18	22 AAF94689	Rho C antisense ph
C 13	13.4	67.0	18	22 AAF94690	Rho C antisense ph
C 14	13.2	66.0	33	14 AAQ36303	HMGCoA reductase t
C 15	13.2	66.0	49	20 AAZ22839	Extreme thermophil
C 16	12.8	64.0	26	14 AAQ52003	B-cell mRNA ribozy
C 17	12.6	63.0	22	19 AAV58750	PCR primer 37803S
C 18	12.6	63.0	27	17 AAT45956	24-Hydroxylase gen
C 19	12.6	63.0	33	19 AAV01423	Human T-cell recep
C 20	12.6	63.0	33	22 AAF98242	C neofomans strai
C 21	12.6	63.0	35	17 AAT45957	24-Hydroxylase gen
C 22	12.6	63.0	37	19 AAV27549	Streptococcus pneu
C 23	12.4	62.0	17	22 AAF61223	Potato dihydro-oro
C 24	12.2	61.0	20	18 AAT48710	Probe for detectin
C 25	12.2	61.0	20	20 AAV73114	Human ras oncogene
C 26	12.2	61.0	20	21 AAC60575	Human fra-1 mRNA a
C 27	12.2	61.0	30	21 AAZ93264	PCR primer h/mPCR
C 28	12.2	61.0	34	20 AAX02783	vector pmwX-BG oli
C 29	12	60.0	15	22 AAF51105	IGF-I oligonucleot
C 30	12	60.0	15	22 AAF51106	IGF-I oligonucleot
C 31	12	60.0	15	22 AAF51107	IGF-I oligonucleot
C 32	12	60.0	15	22 AAF51108	IGF-I oligonucleot
C 33	12	60.0	26	17 AAT69048	Primer for univers
C 34	12	60.0	36	12 AAQ10384	Untranslated regio
C 35	12	60.0	37	20 AAZ08497	Competitor oligonu
C 36	12	60.0	37	20 AAZ08500	G-free binding ele
C 37	12	60.0	39	21 AAZ37033	Calcineurin bindin
C 38	11.8	59.0	20	16 AAQ87917	Human histamine H1
C 39	11.8	59.0	23	21 AAA71892	B. cereus Yisi hom
C 40	11.8	59.0	26	17 AAT69038	Primer for univers
C 41	11.8	59.0	26	17 AAT69046	Primer for univers
C 42	11.8	59.0	26	17 AAT69046	Primer for univers
C 43	11.8	59.0	26	17 AAT69046	Primer for univers
C 44	11.8	59.0	26	17 AAT69015	Primer for univers
C 45	11.8	59.0	26	17 AAT68984	Primer for univers

ALIGNMENTS

RESULT 1

AC65546
ID AAC65546 standard; DNA; 20 BP.

XX AAC65546;

XX 12-FEB-2001 (first entry)

DE Human focal adhesion kinase antisense sequence #12.

Human; focal adhesion kinase; FAK; signal transduction; cancer;
embryonic development disorder; angiogenic disorder; wound healing;
antisense; phosphorothioate; ss.

OS Homo sapiens.

XX US6133031-A.

PD 17-OCT-2000.

PF 19-AUG-1999; 99US-0377310.

PR 19-AUG-1999; 99US-0377310.

XX (ISIS-) ISIS PHARM INC.

XX Monia BP, Gaarde WA;

XX WPI; 2001-006141/01.

PT New antisense compounds for inhibiting focal adhesion kinase
expression, especially useful for inhibiting retinal
neovascularization, or for diagnosing and treating e.g. colon cancer -

PS Claim 15; Column 23; 30pp; English.

CC The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

SQ Sequence 20 BP; 3 A; 3 C; 9 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.35; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggtagggatggtgcccgtca 20
 |||||

Db 1 ggtagggatggtgcccgtca 20
 |||||

RESULT 2

AA655566
 ID AAC65566 standard; DNA; 15 BP.

AC AAC65566;

XX 12-FEB-2001 (first entry)

DE Human focal adhesion kinase antisense sequence #32.

XX Human; focal adhesion kinase; FAK; signal transduction; cancer;
 KW embryonic development disorder; angiogenic disorder; wound healing;
 KW antisense; phosphorothioate; ss.

OS Homo sapiens.

XX US6133031-A.

PN 17-OCT-2000.

PF 19-AUG-1999; 99US-0377310.

PR 19-AUG-1999; 99US-0377310.

XX (ISIS-) ISIS PHARM INC.

PA Monia BP, Gaarde WA;

PI WPI: 2001-006141/01.

DR New antisense compounds for inhibiting focal adhesion kinase
 XX expression, especially useful for inhibiting retinal
 PT neovascularization, or for diagnosing and treating e.g. colon cancer -

PS Example 2; Column 25; 30pp; English.

CC The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

SQ Sequence 15 BP; 2 A; 2 C; 7 G; 4 T; 0 other;

Query Match 75.0%; Score 15; DB 22; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.1e-02; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ttagggatggtgccc 17
 |||||

Db 1 ttagggatggtgccc 15
 |||||

RESULT 3

AAV20799
 ID AAV20799 standard; DNA; 34 BP.

XX AAV20799;

XX 02-JUL-1998 (first entry)

DE Flea serine protease PCR primer SEQ ID NO:58.

XX Flea: Siphonaptera; protease; vaccine; prophylaxis; infestation;
 KW immunoglobulin protease; larvae; host animal; PCR primer; ss.

XX Synthetic.

OS Siphonaptera.

XX WO9740058-A1.

XX 30-OCT-1997.

XX 24-APR-1997; 97WO-US06121.

XX 04-APR-1997; 97US-0042945.

PR 24-APR-1996; 96US-0639075.

PR 15-NOV-1996; 96US-0749699.

XX (HESK-) HESKA CORP.

XX Frank GR, Gaines PJ, Grieve RB, Hunter SW, Rushlow KE;

PI Silver G, Stiegler GL;

XX WPI: 1998-076762/07.

DR New flea protease genes and proteins - used in vaccine compositions
 XX for the prophylaxis and treatment of flea infestation, especially in
 PT cats or dogs

PS Example 23; Page 217; 318pp; English.

CC The present sequence represents a PCR primer for flea serine protease.
 CC The protease, its mimetopes, antibodies (Ab) and inhibitors of the
 CC protein, as well as the DNA encoding the protein, may all be used in
 CC therapeutic compositions to reduce flea protease activity (especially in
 CC immunoglobulin protease) and so reduce flea infestation, especially in
 CC cats or dogs. Alternatively, flea larvae may ingest the faeces of adult
 CC fleas which comprises anti-protease antibodies produced by a host animal
 CC in response to administration of the protein. Therapeutic compositions
 CC may further comprise a compound that reduces haematophagous ectoparasite
 CC burden by a method other than by reducing flea immunoglobulin protease
 CC activity. The novel flea DNA encoding the protein can also be used to
 CC produce recombinant protein, and fragments of it are used as probes and
 CC primers for identification and isolation of related sequences, also as
 CC antisense, triplex-forming agents and ribozymes for inhibition of the
 CC synthesis of the protein. Ab are also useful for screening expression
 CC libraries, to purify the protein and to target cytotoxins to fleas.

XX Sequence 34 BP; 7 A; 11 C; 8 G; 8 T; 0 other;

Query Match 71.0%; Score 14.2; DB 19; Length 34;

Best Local Similarity 84.2%; Pred. No. 2.9e+02;

Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 gttaggatggtgcccgtca 20
 |||||

Db 13 gttaggatggtgcccgtca 31
 |||||

```

RESULT 4
AAQ24636/c
ID AAQ24636 standard; DNA; 24 BP.
XX
AC AAQ24636;
XX
DT 05-NOV-1992 (first entry)
XX
DE Probe Gv003 specific for Gardnerella vaginalis.
XX
KW Dipstick; capture probe; 16S ribosomal RNA; vaginal infection; ss.
XX
OS Synthetic.
XX
PN W09207096-A.
PD 30-APR-1992.
XX
PF 21-OCT-1991; 91WO-US07763.
XX
PR 19-OCT-1990; 90US-0600334.
XX
PA (MICR-) MICROPROBE CORP.
PI Adams TH, Sheiness DK;
XX
DR WPI; 1992-167174/20.
XX
PT Detecting bacterial vaginitis infections - by contacting formed
PT lysate with hybridisation soln. and dipstick comprising
PT non-porous solid support and determining signal associated with
PT beads
XX
PS Disclosure; Page 39; 62pp; English.
XX
CC This oligonucleotide is one example of a probe suitable for
CC diagnosing the presence of G.vaginalis. The preferred target for
CC such probes is 16S rRNA, present in several thousand copies per cell.
CC Oligonucleotides complementary to sequences in the G.vaginalis
CC genome or in G.vaginalis plasmids may also be used.
CC See also AAQ24637-Q24647.
XX
SQ Sequence 24 BP; 7 A; 7 C; 6 G; 4 T; 0 other;

Query Match 69.0%; Score 13.8; DB 13; Length 24;
Best Local Similarity 88.2%; Pred. No. 4.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ttagggtggtgccgtc 19
DB 18 TTTGGGATGGAGCCGTC 2

RESULT 5
AAQ52752/c
ID AAQ52752 standard; DNA; 24 BP.
XX
AC AAQ52752;
XX
DT 24-JUN-1994 (first entry)
XX
DE Sequence of probe Gv003 for the detection of Gardnerella vaginalis.
XX
KW Probe; vaginal fluid microorganism; oligonucleotide; gram positive;
KW ss.
XX
OS Synthetic.
XX
PN W09324659-A.
XX
PD 09-DEC-1993.
XX

Query Match 69.0%; Score 13.8; DB 13; Length 24;
Best Local Similarity 88.2%; Pred. No. 4.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ttagggtggtgccgtc 19
DB 18 TTTGGGATGGAGCCGTC 2

RESULT 6
AAT89823/c
ID AAT89823 standard; DNA; 24 BP.
XX
AC AAT89823;
XX
DT 27-FEB-1998 (first entry)
XX
DE Oligonucleotide probe Gv003 specific for Gardnerella vaginalis.
XX
KW Probe; hybridise; ribosomal RNA; vaginitis; infection; diagnostic;
KW ribosomal RNA; ss.
XX
OS Synthetic.
OS Gardnerella vaginalis.
XX
PN US5654418-A.
XX
PD 05-AUG-1997.
XX
PF 02-JUN-1995; 95US-0460344.
XX
PR 08-OCT-1993; 93US-0133598.
PR 19-OCT-1990; 90US-0600334.
PR 29-MAY-1992; 92US-0896094.
XX
PA (BECT ) BECTON DICKINSON CO.
XX
PI Britschgi TB, Cangelosi GA, Sheiness DK;
XX
DR WPI; 1997-401906/37.
XX
Oligo:nucleotide probes for detecting microorganisms in vaginal
PT samples - allows selective detection of vaginitis-causing
PT microorganisms
XX
PS Claim 1; Column 32; 40pp; English.
XX

```

```

PF 28-MAY-1993; 93WO-US05085.
XX
PR 29-MAY-1992; 92US-0896094.
XX
PA (MICR-) MICROPROBE CORP.
XX
PI Adams TH, Britschgi TB, Cangelosi GA, Sheiness DK;
PI Stamm MR;
XX
DR WPI; 1993-405847/50.
XX
PT Detection of microorganisms, partic. in vaginal fluid samples -
PT by lysis of cells to release nucleic acid and hybridisation
PT assays using probes
XX
PS Claim 9; Page 42; 78pp; English.
XX
CC The inventors claim a compsn. of at least one oligo probe for the
CC detection of at least one microorganism found in vaginal samples.
CC The sequence below is one such probe. The methods and compsns. can
CC be used for the detection of microorganisms, partic. in vaginal
CC fluid samples.
XX
SQ Sequence 24 BP; 7 A; 7 C; 6 G; 4 T; 0 other;

Query Match 69.0%; Score 13.8; DB 14; Length 24;
Best Local Similarity 88.2%; Pred. No. 4.5e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ttagggtggtgccgtc 19
DB 18 TTTGGGATGGAGCCGTC 2

RESULT 6
AAT89823/c
ID AAT89823 standard; DNA; 24 BP.
XX
AC AAT89823;
XX
DT 27-FEB-1998 (first entry)
XX
DE Oligonucleotide probe Gv003 specific for Gardnerella vaginalis.
XX
KW Probe; hybridise; ribosomal RNA; vaginitis; infection; diagnostic;
KW ribosomal RNA; ss.
XX
OS Synthetic.
OS Gardnerella vaginalis.
XX
PN US5654418-A.
XX
PD 05-AUG-1997.
XX
PF 02-JUN-1995; 95US-0460344.
XX
PR 08-OCT-1993; 93US-0133598.
PR 19-OCT-1990; 90US-0600334.
PR 29-MAY-1992; 92US-0896094.
XX
PA (BECT ) BECTON DICKINSON CO.
XX
PI Britschgi TB, Cangelosi GA, Sheiness DK;
XX
DR WPI; 1997-401906/37.
XX
Oligo:nucleotide probes for detecting microorganisms in vaginal
PT samples - allows selective detection of vaginitis-causing
PT microorganisms
XX
PS Claim 1; Column 32; 40pp; English.
XX

```


Db 11 ggtagggatgtagcg 27
|||||

RESULT 9

AAF70762
ID AAF70762 standard; DNA; 30 BP.

XX
AC AAF70762;

XX
DT 20-APR-2001 (first entry)

XX
DE Thrombin high affinity ligand #9.

XX
KW Ligand; basic fibroblast growth factor; bFGF; gene therapy; vascular;
KW atherosclerosis; angioplasty; stability; ss.

XX
OS Unidentified.

XX
PN US6177557-B1.

XX
PD 23-JAN-2001.

XX
PF 05-AUG-1996; 96US-0687421.

XX
PR 11-JUN-1990; 90US-0536428.

XX
PR 10-JUN-1991; 91US-0714131.

XX
PR 06-NOV-1992; 92US-0973333.

XX
PR 10-FEB-1994; 94US-0195005.

XX
PR 28-MAR-1994; 94US-0219012.

XX
PA (NEXS-) NEXSTAR PHARM INC.

XX
PI Janjic N, Gold L, Tasset D;

XX
XX WPI; 2001-158583/16.

XX
PT Novel nucleic acid ligands to basic fibroblast growth factor that are
PT useful as inhibitors of basic fibroblast growth factors and 2'-amino
PT modified RNA ligands, exhibit increased in vivo stability -

XX
XX Example 19; Column 57-58; 153pp; English.

XX
CC The present invention relates to a purified and isolated non-naturally
CC occurring DNA ligands to basic fibroblast growth factor (bFGF).
CC The ligands are useful as part of gene therapy treatments and
CC for diagnosing pathogenesis of vascular diseases including
CC initiation and progression of atherosclerosis, acute coronary
CC syndromes, vein graft disease and restenosis following coronary
CC angioplasty. The ligands have improved stability in vivo.

XX
SQ Sequence 30 BP; 4 A; 1 C; 15 G; 10 T; 0 other;

Query Match 69.0%; Score 13.8; DB 22; Length 30;

Best Local Similarity 88.2%; Pred. No. 4.6e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggtagggatgtagcg 17
|||||

Db 11 ggtagggatgtagcg 27

RESULT 10

AAZ40968/c
ID AAZ40968 standard; DNA; 18 BP.

XX
AC AAZ40968;

XX
DT 26-JAN-2000 (first entry)

XX
DE Human RhoC phosphorothioate antisense oligonucleotide SEQ ID NO:120.

XX

KW Identification; genetic target; gene modulation; human; probe;
KW antisense oligonucleotide; phosphorothioate; PCR primer;
KW nucleotide sequence-based technology; antisense drug discovery;
KW target validation; ss.

XX
OS Synthetic.

OS
XX Homo sapiens.

XX
PN WO9953101-A1.

XX
PD 21-OCT-1999.

XX
PF 13-APR-1999; 99WO-US08268.

XX
PR 13-APR-1998; 98US-0081483.

XX
PR 28-APR-1998; 98US-0067638.

XX
PA (ISIS-) ISIS PHARM INC.

XX
PI Cowsert LM, Baker BF, McNeil J, Freier SM, Sasmor HM, Brooks DG;
PI Ohasi C, Wyatt JR, Borchers AH, Vickers TA;

XX
DR WPI; 1999-620446/53.

XX
PT Identifying compounds which modulate expression of nucleic acids, used
PT to provide compounds having defined physical, chemical or bioactive
PT properties, e.g. antisense activity -

XX
PS Example 18; Page 96; 264pp; English.

XX
CC A method has been developed of defining a set of compounds that modulate
CC the expression of a target nucleic acid (tNA) sequence via binding of
CC the compounds with the tNA sequence. The method comprises generating a
CC library of virtual compounds in silico according to defined criteria,
CC and evaluating in silico the binding of the virtual compounds with the
CC tNA according to defined criteria. Also described are: (1) a method of
CC defining a set of oligonucleotides (ONs) that modulate the expression of
CC a tNA sequence via binding of the ONs with the tNA sequence comprising
CC generating a library of virtual compounds in silico according to defined
CC criteria, and evaluating in silico the binding of the virtual ONs with
CC the tNA according to defined criteria; and (2) a method of defining a
CC set of compounds that modulate the expression of a tNA sequence via
CC binding of the compounds with the tNA. The methods can be used for the
CC generation and identification of synthetic compounds having defined
CC physical, chemical or bioactive properties. Information gathered from
CC assays of such compounds is used to identify nucleic acid sequences that
CC are tractable to a variety of nucleotide sequence-based technologies,
CC e.g. antisense drug discovery and target validation. AAZ40852 to
CC AAZ41220, and AAY52701 to AAY52706, represent sequences used in the
CC exemplification of the present invention.

XX
SQ Sequence 18 BP; 5 A; 8 C; 3 G; 2 T; 0 other;

Query Match 67.0%; Score 13.4; DB 20; Length 18;

Best Local Similarity 93.3%; Pred. No. 7e+02;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gttagggatggtgcc 16
|||

Db 15 GTTGGGGATGGTGC 1

RESULT 11

AAZ40969/c
ID AAZ40969 standard; DNA; 18 BP.

XX
AC AAZ40969;

XX
DT 26-JAN-2000 (first entry)

XX
DE Human RhoC phosphorothioate antisense oligonucleotide SEQ ID NO:121.

XX

KW Identification; genetic target; gene modulation; human; probe;
 KW antisense oligonucleotide; phosphorothioate; PCR primer;
 KW nucleotide sequence-based technology; antisense drug discovery;
 KW target validation; ss.

XX Synthetic.

OS Homo sapiens.

XX WO9953101-A1.

PN 21-OCT-1999.

XX 13-APR-1999; 99WO-US08268.

XX 13-APR-1998; 98US-0081483.

XX 28-APR-1998; 98US-0067638.

XX (ISIS-) ISIS PHARM INC.

PI Cowsett LM, Baker BF, McNeil J, Freier SM, Sasmor HM, Brooks DG;

PI Ohasi C, Wyatt JR, Borchers AH, Vickers TA;

XX WPI; 1999-620446/53.

XX Identifying compounds which modulate expression of nucleic acids, used

PT to provide compounds having defined physical, chemical or bioactive

PT properties, e.g. antisense activity

XX Example 18; Page 96; 264pp; English.

XX A method has been developed of defining a set of compounds that modulate

CC the expression of a target nucleic acid (tNA) sequence via binding of

CC the compounds with the tNA sequence. The method comprises generating a

CC library of virtual compounds in silico according to defined criteria,

CC and evaluating in silico the binding of the virtual compounds with the

CC tNA according to defined criteria. Also described are: (1) a method of

CC defining a set of oligonucleotides (ONS) that modulate the expression of

CC a tNA sequence via binding of the ONS with the tNA sequence comprising

CC generating a library of virtual compounds in silico according to defined

CC criteria, and evaluating in silico the binding of the virtual ONS with

CC the tNA according to defined criteria; and (2) a method of defining a

CC set of compounds that modulate the expression of a tNA sequence via

CC binding of the compounds with the tNA. The methods can be used for the

CC generation and identification of synthetic compounds having defined

CC physical, chemical or bioactive properties. Information gathered from

CC assays of such compounds is used to identify nucleic acid sequences that

CC are tractable to a variety of nucleotide sequence-based technologies,

CC e.g. antisense drug discovery and target validation. AAZ40852 to

CC AAZ41220, and AAZ52701 to AAZ52706, represent sequences used in the

CC exemplification of the present invention.

XX Sequence 18 BP; 6 A; 8 C; 3 G; 1 T; 0 other;

SQ

Query Match

Best Local Similarity 67.0%; Score 13.4; DB 20; Length 18;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gttaggagtggtgcc 16

||| |||||

Db 16 GTTGGGATGGTGCC 2

RESULT 12

AAAF94689/c

ID AAFA94689 standard; DNA; 18 BP.

XX AAF94689;

AC AAF94689;

XX 23-MAY-2001 (first entry)

DT Rho C antisense phosphorothioate oligonucleotide SEQ ID 113.

XX

XX

XX

XX

XX

XX

XX

XX

KW Rho; GTP binding protein; phosphorothioate antisense oligonucleotide;
 KW RhoA; RhoB; RhoC; RhoG; Rac 1; cdc42; hyperproliferative condition;
 KW cancer; wound healing; clotting; ischaemia; reperfusion; reoxygenation;
 KW ss.

XX Homo sapiens.

OS WO200115739-A1.

XX 08-MAR-2001.

XX 18-AUG-2000; 2000WO-US22808.

XX 31-AUG-1999; 99US-0387341.

XX (ISIS-) ISIS PHARM INC.

XX Roberts ML, Cowsett LM;

PI WPI; 2001-191677/19.

PI An antisense compound targeted to a nucleic acid molecule encoding a

PT member of the human Rho family of small GTP binding proteins useful for

PT treating e.g. cancer and ischaemia

XX Example 16; Page 73; 156pp; English.

XX This invention relates to an antisense compound targeted to a nucleic

CC acid molecule encoding a member of the human Rho family of small GTP

CC binding proteins, where the antisense compound inhibits the expression of

CC the member of the human Rho family. The invention includes antisense

CC oligonucleotides AAF94580 - AAF94637 which target a RhoA nucleotide

CC sequence, AAF94645 - AAF94684 which target a RhoB nucleotide sequence,

CC AAF94686 - AAF94725 which target a RhoC nucleotide sequence, AAF94727 -

CC AAF94766 which target RhoG nucleotide sequence, AAF94769 - AAF94790 which

CC target a Rac 1 nucleotide sequence and AAF94795 - AAF94809 which target

CC cdc42 nucleotide sequence. The antisense compound is useful for treating

CC hyperproliferative conditions, especially cancer, abnormal wound healing

CC or clotting conditions and ischaemia/reperfusion or reoxygenation injury.

CC The compound may also be used to diagnose the above conditions.

XX Sequence 18 BP; 5 A; 8 C; 3 G; 2 T; 0 other;

SQ

Query Match

Best Local Similarity 67.0%; Score 13.4; DB 22; Length 18;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gttaggagtggtgcc 16

||| |||||

Db 15 GTTGGGATGGTGCC 1

RESULT 13

AAAF94690/c

ID AAFA94690 standard; DNA; 18 BP.

XX AAF94690;

AC AAF94690;

XX 23-MAY-2001 (first entry)

DT Rho C antisense phosphorothioate oligonucleotide SEQ ID 114.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX 18-AUG-2000; 2000WO-US22808.
 XX 31-AUG-1999; 99US-0387341.
 XX (ISIS-) ISIS PHARM INC.
 XX Roberts ML, Cowseert LM;
 XX WPI; 2001-191677/19.
 XX An antisense compound targeted to a nucleic acid molecule encoding a
 PT member of the human Rho family of small GTP binding proteins useful for
 PT treating e.g. cancer and ischaemia -
 XX Example 16; Page 73; 156pp; English.
 XX This invention relates to an antisense compound targeted to a nucleic
 CC acid molecule encoding a member of the human Rho family of small GTP
 CC binding proteins, where the antisense compound inhibits the expression of
 CC the member of the human Rho family. The invention includes antisense
 CC oligonucleotides AAF94580 - AAF94637 which target a RhoA nucleotide
 CC sequence, AAF94645 - AAF94684 which target a RhoB nucleotide sequence,
 CC AAF94686 - AAF94725 which target a RhoC nucleotide sequence, AAF94727 -
 CC AAF94766 which target RhoG nucleotide sequence, AAF94769 - AAF94790 which
 CC target a Rac 1 nucleotide sequence and AAF94795 - AAF94809 which target
 CC cdc42 nucleotide sequence. The antisense compound is useful for treating
 CC hyperproliferative conditions, especially cancer, abnormal wound healing
 CC or clotting conditions and ischaemia/reperfusion or reoxygenation injury.
 CC The compound may also be used to diagnose the above conditions.
 XX Sequence 18 BP; 6 A; 8 C; 3 G; 1 T; 0 other;

Query Match 67.0%; Score 13.4; DB 22; Length 18;
 Best Local Similarity 93.3%; Pred. No. 7e+02;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 gttagggtggtgccc 16
 ||| |||||
 Db 16 GTTGGGGATGGTGC 2

RESULT 14
 AAQ36303
 ID AAQ36303 standard; DNA; 33 BP.
 XX
 AC AAQ36303;
 XX
 DT 07-JUN-1993 (first entry)
 XX
 DE HMGCoA reductase target sequence.
 XX
 KW Cholesterol; repressor protein; transcription; agonist; cellular;
 KW triplex; target; duplex; ss.
 XX
 OS Synthetic.
 XX
 PN US5176996-A.
 XX
 PD 05-JAN-1993.
 XX
 PF 20-DEC-1988; 88US-0287359.
 XX
 PR 20-DEC-1988; 88US-0287359.
 PR 22-DEC-1989; 89US-0453532.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Hogan ME, Kessler DJ;
 XX
 DR WPI; 1993-035718/04.
 XX

PT Synthetic oligo-nucleotide(s), prodn. useful e.g. for HIV-1
 PT inhibition - which bind to target sequence in duplex DNA forming
 PT colinear triplex by binding to major groove
 XX Example 9; Column 28; 29pp; English.
 XX HMGCoA reductase is the enzyme which defines the rate limiting step
 CC in cholesterol biosynthesis. Expression of the gene may be
 CC prevented by the formation of a triplex between the duplex target DNA
 CC sequence and an anti parallel or parallel synthetic oligonucleotide.
 CC A suitable target sequence is the binding site for a repressor
 CC protein which appears to mediate end-prod. inhibition of transcription
 CC by cholesterol, from base -167 to -135 of HMGCoA reductase gene.
 CC Oligonucleotides targetted against this sequence will repress
 CC HMGCoA expression, and act as agonists of the cellular repressor.
 CC See also AAQ36219-362.
 XX Sequence 33 BP; 3 A; 10 C; 13 G; 7 T; 0 other;

Query Match 66.0%; Score 13.2; DB 14; Length 33;
 Best Local Similarity 83.3%; Pred. No. 9.3e+02;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggttagggatggtgcgct 18
 ||| |||||
 Db 1 ggtgagagatggtgcggt 18

RESULT 15
 AAZ22839/c
 ID AAZ22839 standard; DNA; 49 BP.
 XX
 AC AAZ22839;
 XX
 DT 06-DEC-1999 (first entry)
 XX
 DE Extreme thermophilic mismatch cleavage enzyme substrate oligo #4.
 XX
 KW Thermophilic; mismatch; cleavage; PCR; amplification; Taq polymerase;
 KW mutation; detection; pathogen; cancer; probe; ss.
 XX
 OS Synthetic.
 XX
 PN WO9942595-A1.
 XX
 PD 26-AUG-1999.
 XX
 PF 19-FEB-1999; 99WO-US03274.
 XX
 PR 19-FEB-1998; 98US-0075194.
 XX
 PA (TREV-) TREVIGEN INC.
 XX
 PI Chirikjian JG, Bazar LS, George AL;
 XX
 DR WPI; 1999-571571/48.
 XX
 PT New isolated extreme thermophilic enzyme, which cleaves at a mismatch
 PT formed between two polynucleotides in a duplex -
 XX Example 4; Fig 2; 41pp; English.

XX This sequence represents oligonucleotide #4 which was annealed to
 CC oligonucleotides #1, #2 and #3 (AAZ22836-22838) to generate a four-way
 CC junction used to assay for any resolvase activity in an extreme
 CC thermophilic enzyme composition from Thermotoga maritima. Extreme
 CC thermophilic enzymes, such as TM-Endo V (from Thermotoga maritima)
 CC cleave at a mismatch formed between two polynucleotides in a duplex. It
 CC does not exhibit resolvase activity and does not require a GATC sequence
 CC to effect cleavage. TM-Endo V cleaves C/C, G/G, T/C, A/C, A/A and T/T
 CC mismatch. The enzyme does not cleave a perfectly complementary
 CC oligonucleotide pair (no mismatch), a T/G mismatch or a 3 base bubble. It

CC can be used to detect mutations in a target nucleotide or to detect the
CC presence of a non-mutated sequence. In particular, the enzyme can be used
CC for detecting mutations in HIV, sickle cell anaemia hypoxanthine
CC transferase gene, and in genes associated with cancer e.g., p53 and
CC BRCA1. TM-Endo V can be used to cleave mismatches that are created during
CC PCR amplification. The extreme thermophilic DNA polymerase, Taq DNA
CC polymerase, incorporates approximately one incorrect base per 500 bases,
CC which is unacceptable if DNA greater than 500 bp is to be amplified.
CC TM-Endo V cleaves at the mismatch, either on the template or the
CC daughter strand, preventing the cleaved strand from undergoing any
CC further amplification. A chimera of an extreme thermophilic DNA
CC polymerase and TM-Endo V can accomplish the dual roles of PCR plus
CC correction of erroneous base incorporation.
xx
SQ Sequence 49 BP; 13 A; 15 C; 12 G; 9 T; 0 other;

Query Match 66.0%; Score 13.2; DB 20; Length 49;
Best Local Similarity 83.3%; Pred. No. 9.6e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gtaggagtggtccgctc 19
 ||||| |||
Db 31 GTTAGGCTCGTCTGTC 14

Search completed: October 2, 2001, 16:18:42
Job time: 15486 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:00 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-14

Perfect score: 20

Sequence: 1 ggttagggatggtgccgtca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_est4.*
- 5: gb_est5.*
- 6: gb_est6.*
- 7: gb_est7.*
- 8: gb_est8.*
- 9: gb_est9.*
- 10: gb_est10.*
- 11: gb_est11.*
- 12: gb_est12.*
- 13: gb_est13.*
- 14: gb_est14.*
- 15: gb_est15.*
- 16: gb_est16.*
- 17: gb_est17.*
- 18: gb_est18.*
- 19: gb_est19.*
- 20: gb_est20.*
- 21: gb_est21.*
- 22: gb_est22.*
- 23: gb_est23.*
- 24: gb_est24.*
- 25: gb_est25.*
- 26: gb_est26.*
- 27: gb_est27.*
- 28: gb_est28.*
- 29: gb_est29.*
- 30: gb_est30.*
- 31: gb_est31.*
- 32: gb_est32.*
- 33: gb_est33.*
- 34: gb_est34.*
- 35: gb_est35.*
- 36: gb_est36.*
- 37: gb_est37.*
- 38: gb_est38.*
- 39: gb_est39.*
- 40: gb_est40.*
- 41: gb_est41.*
- 42: gb_est42.*
- 43: gb_est43.*
- 44: gb_est44.*
- 45: gb_est45.*
- 46: gb_est46.*
- 47: gb_est47.*

- 44: em_esthum10.*
- 45: em_esthum11.*
- 46: em_esthum12.*
- 47: em_esthum13.*
- 48: em_esthum14.*
- 49: em_esthum15.*
- 50: em_esthum16.*
- 51: em_esthum17.*
- 52: em_esthum18.*
- 53: em_esthum19.*
- 54: em_esthum20.*
- 55: em_esthum21.*
- 56: em_esthum22.*
- 57: em_esthum23.*
- 58: em_esthum24.*
- 59: em_esthum25.*
- 60: em_esthum26.*
- 61: em_esthum27.*
- 62: em_esthum28.*
- 63: em_estin1.*
- 64: em_estin2.*
- 65: em_estin3.*
- 66: em_estin4.*
- 67: em_estin5.*
- 68: em_estom1.*
- 69: em_estom2.*
- 70: em_estov1.*
- 71: em_estov2.*
- 72: em_estpl1.*
- 73: em_estpl2.*
- 74: em_estpl3.*
- 75: em_estpl4.*
- 76: em_estpl5.*
- 77: em_estpl6.*
- 78: em_estpl7.*
- 79: em_estpl8.*
- 80: em_estpl9.*
- 81: em_estpl10.*
- 82: em_estro1.*
- 83: em_estro2.*
- 84: em_estro3.*
- 85: em_estro4.*
- 86: em_estro5.*
- 87: em_estro6.*
- 88: em_estro7.*
- 89: em_estro8.*
- 90: em_estro9.*
- 91: em_estro10.*
- 92: em_estro11.*
- 93: em_estro12.*
- 94: em_estro13.*
- 95: em_estro14.*
- 96: em_estro15.*
- 97: em_estro16.*
- 98: em_estro17.*
- 99: em_estro18.*
- 100: em_estro19.*
- 101: em_estro20.*
- 102: gb_est25.*
- 103: gb_est26.*
- 104: gb_est27.*
- 105: gb_est28.*
- 106: gb_est29.*
- 107: gb_est30.*
- 108: gb_est31.*
- 109: gb_est32.*
- 110: gb_est33.*
- 111: gb_est34.*
- 112: gb_est35.*
- 113: gb_est36.*
- 114: gb_est37.*
- 115: gb_est38.*
- 116: gb_est39.*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

TITLE and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: dunn@genetics.utah.edu
 Insert length: 10000 Std Error: 0.00
 Plate: 0402 row: H column: 06
 Seq primer: CACACAGAAACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 29.

FEATURES

source
 1. .29
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0402H06"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource.
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (g14732114/gb/AP129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 5 a 7 c 10 g 7 t
 ORIGIN

Query Match 65.0%; Score 13; DB 246; Length 29;
 Best Local Similarity 100.0%; Pred. No. 2.6e+04;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 gatgagccatca 20
 |||||
 Db 2 GATGGCCGCTCA 14

RESULT 3
 AA009110/c
 LOCUS
 DEFINITION AA009110 45 bp mRNA EST 26-JUL-1996
 mg92g09.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
 clone IMAGE:440512 5' similar to gb:X05021 Murine mRNA with
 homology to yeast L29 ribosomal protein (MOUSE); mRNA sequence.
 AA009110
 ACCESSION AA009110.1 GI:1465176
 VERSION
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 1 (bases 1 to 45)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE
 JOURNAL
 COMMENT

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Theising, B., Wylie, T., Lannon, G., Soares, B., Wilson, R. and
 Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:265848

Trace considered overall poor quality

Seq primer: Etpriimer
 High quality sequence stop: 1.

FEATURES

source
 1. .45
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:440512"
 /clone_lib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TCTTACCAATCGAAGTGGAGCGCGGAAATTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2]; Double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."

BASE COUNT 15 a 14 c 8 g 8 t
 ORIGIN

Query Match 64.0%; Score 12.8; DB 1; Length 45;

Best Local Similarity 87.5%; Pred. No. 3.4e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggttaggagtggtgcc 16
 |||||
 Db 43 GGTTTGGATGGTGC 28

RESULT 4
 AZ607381

LOCUS
 DEFINITION IM0429K17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0429K17 R, DNA sequence.

ACCESSION AZ607381
 VERSION
 KEYWORDS GSS.

SOURCE
 ORGANISM

Mus musculus
 house mouse.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL
COMMENT

plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0429 row: K column: 17
 Seq primer: CACACGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 27.

FEATURES

Location/Qualifiers
 1..27
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0429K17"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gii14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT
ORIGIN

5 a 9 c 6 g 7 t

Query Match 61.0%; Score 12.2; DB 246; Length 27;
 Best Local Similarity 82.4%; Pred. No. 6.4e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ttagggatgggtgcgc 19
 ||| ||||| ||| ||
 Db 3 TTATGGATGGGCCATC 19

RESULT 5

AA501145 49 bp mRNA EST 01-JUL-1997
 LOCUS vH3d08.r1 Knowles Solter mouse inner cell mass Mus musculus cDNA
 DEFINITION clone IMAGE:892623 5' similar to WP:B0491.1 CE02103 ;, mRNA
 sequence.

ACCESSION AA501145.1 GI:2236112
 VERSION AA501145
 KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 49)
 AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

TITLE
JOURNAL
COMMENT

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LINL : contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:520583

Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 High quality sequence stop: 1.

FEATURES
source

Location/Qualifiers
 1..49
 /organism="Mus musculus"
 /strain="B6D2 F1/J"
 /db_xref="taxon:10090"
 /clone="IMAGE:892623"
 /clone_lib="Knowles Solter mouse inner cell mass"
 /dev_stage="embryo (post-implantation)"
 /lab_host="DH10B"
 /note="Vector: pBluescript SK+; Site1: XbaI; Site2: XhoI
 : Cloned unidirectionally from mRNA prepared from
 primitive streak embryonic tissue. Primer: Oligo dT.
 cDNAs were cloned into the XbaI/XhoI sites of pBluescript
 SK+ (Stratagene) using commercial linkers (NEB). Average
 insert size: 0.5 kb."

BASE COUNT 14 a 9 c 13 g 13 t
 ORIGIN

Query Match 61.0%; Score 12.2; DB 8; Length 49;
 Best Local Similarity 82.4%; Pred. No. 6.6e+04;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggttagggatggtgcgc 17
 ||||| ||||| ||| ||
 Db 21 GGTAGAGATGATGCGC 37

RESULT 6
A2434508/c

LOCUS A2434508 23 bp DNA GSS 03-OCT-2000
 DEFINITION 1M0220K22R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0220K22 R, DNA sequence.

ACCESSION A2434508.1 GI:10559521
 VERSION A2434508
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 23)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
 Plate: 0220 row: K column: 22
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 23.
 Location/Qualifiers

FEATURES

source
 1. .23
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM0220R22"
 /clone_lib="Mouse 10kb plasmid UUGCLM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 9 a 8 c 4 g 2 t
 ORIGIN

Query Match 60.0%; Score 12; DB 243; Length 23;
 Best Local Similarity 75.0%; Pred. No. 8e+04;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 ggtagggatgggtgcctca 20
 ||||| | ||||| |||
 Db 20 GGTTCGTGCTGCTGCTCA 1

RESULT 7

AZ769994
 LOCUS AZ769994 47 bp DNA GSS 16-FEB-2001
 DEFINITION IM0571A14F Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCLM0571A14 F, DNA sequence.
 ACCESSION AZ769994
 VERSION AZ769994.1 GI:12890717
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 47)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0571 row: A column: 14
 Seq primer: CGTTGTAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 47.
 Location/Qualifiers

FEATURES

source
 1. .47
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM0571A14"
 /clone_lib="Mouse 10kb plasmid UUGCLM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 5 a 8 c 25 g 9 t
 ORIGIN

Query Match 60.0%; Score 12; DB 249; Length 47;
 Best Local Similarity 100.0%; Pred. No. 8.3e+04;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 agggatgggtgcc 16
 ||||| ||||| |||
 Db 21 AGGATGGTGCC 32

RESULT 8

AZ760007
 LOCUS AZ760007 37 bp DNA GSS 16-FEB-2001
 DEFINITION IM0553C23F Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCLM0553C23 F, DNA sequence.
 ACCESSION AZ760007
 VERSION AZ760007.1 GI:12867379
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 37)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606

```

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0553 row: C column: 23
Seq primer: CGTGTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 37.
Location/Qualifiers
1. 37
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0553C23"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="F. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (91473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
11 a 6 c 7 g 13 t
BASE COUNT
ORIGIN

Query Match 59.0%; Score 11.8; DB 249; Length 37;
Best Local Similarity 86.7%; Pred. No. 1e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ggtagggatggtgc 15
|||||
Db 5 GGTGTGGATGTC 19

RESULT 9
AUI03569/c
LOCUS
DEFINITION AUI03569 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
CAS00968, mRNA sequence.
ACCESSION AUI03569
VERSION AUI03569.1 GI:13553090
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 50)
Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isoqai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.
Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries
Unpublished (2001)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano

,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
source
1. 50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CAS00968"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 4 a 14 c 22 g 10 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 1e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 gggatgggtccgtca 20
|||||
Db 39 GGGATGCTGCCGACA 25

RESULT 10
A2345513/c
LOCUS
DEFINITION A2345513 20 bp DNA GSS 29-SEP-2000
clone UUGC1M0080J04 F, DNA sequence.
ACCESSION A2345513
VERSION A2345513.1 GI:10424750
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 20)
Dunn,D., Aoyagi,A., Barber,M., Becorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0080 row: J column: 04
Seq primer: CGTCTAAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1. 20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0080J04"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="F. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to

```


AZ399400/c
 LOCUS 1M0165C13F 21 bp DNA GSS 03-OCT-2000
 DEFINITION clone UUGC1M0165C13 F, DNA sequence.
 ACCESSION AZ399400
 VERSION AZ399400.1 GI:10514474
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 21)
 REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 M., Rose,M., Rose,R., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 and Wright,D., Weiss,R., Stokes,R., Tingey,A., von Niederhausern,A.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0165 row: C column: 13
 Seq primer: CGTGTAAACGACGGCCAGT
 Class: plasmid ends
 High quality sequence stop: 21.
 Location/Qualifiers
 1..21
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0165C13"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi14732114|gbAF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 5 a 11 c 0 g 5 t
 ORIGIN
 Query Match 57.0%; Score 11.4; DB 243; Length 21;
 Best Local Similarity 92.3%; Pred. No. 1.6e+05;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 gttaggatggtg 14
 |||||
 Db 14 GTTAGGATGATG 2

RESULT 14
 AZ786429
 LOCUS 2M0031L23R 21 bp DNA GSS 16-FEB-2001
 DEFINITION clone UUGC2M0031L23 R, DNA sequence.
 ACCESSION AZ786429
 VERSION AZ786429.1 GI:12924180
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 21)
 REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 M., Rose,M., Rose,R., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 and Wright,D., Weiss,R., Stokes,R., Tingey,A., von Niederhausern,A.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0031 row: L column: 23
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 21.
 Location/Qualifiers
 1..21
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0031L23"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi14732114|gbAF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 2 a 1 c 13 g 5 t
 ORIGIN
 Query Match 57.0%; Score 11.4; DB 249; Length 21;
 Best Local Similarity 92.3%; Pred. No. 1.6e+05;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 2 gttaggatggtg 14
 |||||
 Db 8 GTTGGGATGGTG 20

RESULT 15

AZ822907 38 bp DNA GSS 20-FEB-2001
LOCUS AZ822907 Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0096M03 R, DNA sequence.

ACCESSION AZ822907

VERSION AZ822907.1 GI:12992815

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 38)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

Muscle whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert length: 10000 Std Error: 0.00

Plate: 0096 row: M column: 03

Seq primer: CACACAGGAACACGCTATGACC

Class: plasmid ends

High quality sequence stop: 38.

FEATURES

source

1..38
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0096M03"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 10 a 8 c 11 g 9 t

ORIGIN

Query Match 57.0%; Score 11.4; DB 250; Length 38;
Best Local Similarity 92.3%; Pred. No. 1.6e+05;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ggttagggatggt 13

|||||

Db 20 GGTGCGGGTGGT 32

Search completed: October 2, 2001, 15:01:02
Job time: 10831 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:50 ; Search time 417.38 Seconds
(without alignments)
9.071 Million cell updates/sec

Title: US-09-757-100B-14

Perfect score: 20

Sequence: 1 ggttagggatggtgcgtca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCITUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	3	US-09-377-310-14
2	15	75.0	15	3	US-09-377-310-34
3	14.2	71.0	34	4	US-09-004-731-58
4	14.2	71.0	34	4	US-08-749-699-58
5	13.8	69.0	24	1	US-08-460-344-18
6	13.8	69.0	24	1	US-08-133-598A-18
7	13.8	69.0	24	1	US-08-886-999-18
8	13.8	69.0	24	5	PCIT-US93-05085-18
9	13.8	69.0	30	1	US-08-219-012-36
10	13.8	69.0	30	4	US-08-687-421-224
11	13.4	67.0	18	3	US-09-156-807-11
12	13.4	67.0	18	3	US-09-156-807-12
13	12.6	63.0	37	3	US-08-961-083-339
14	12.6	63.0	50	6	5240845-45
15	12.2	61.0	20	3	US-09-418-641-75
16	12	60.0	26	1	US-08-429-121B-87
17	12	60.0	26	2	US-09-003-067-87
18	11.8	59.0	26	1	US-08-429-121B-5
19	11.8	59.0	26	1	US-08-429-121B-27
20	11.8	59.0	26	1	US-08-429-121B-55
21	11.8	59.0	26	1	US-08-429-121B-65
22	11.8	59.0	26	1	US-08-429-121B-67
23	11.8	59.0	26	1	US-08-429-121B-81
24	11.8	59.0	26	1	US-08-429-121B-83
25	11.8	59.0	26	1	US-08-429-121B-85
26	11.8	59.0	26	2	US-09-003-067-5
27	11.8	59.0	26	2	US-09-003-067-27

c 28	11.8	59.0	26	2	US-09-003-067-55	Sequence 55, Appl
c 29	11.8	59.0	26	2	US-09-003-067-65	Sequence 65, Appl
c 30	11.8	59.0	26	2	US-09-003-067-67	Sequence 67, Appl
c 31	11.8	59.0	26	2	US-09-003-067-81	Sequence 81, Appl
c 32	11.8	59.0	26	2	US-09-003-067-83	Sequence 83, Appl
c 33	11.8	59.0	26	2	US-09-003-067-85	Sequence 85, Appl
c 34	11.8	59.0	26	2	US-08-859-998-933	Sequence 933, Appl
c 35	11.8	59.0	30	1	US-08-219-012-50	Sequence 50, Appl
c 36	11.8	59.0	30	4	US-08-687-421-238	Sequence 238, Appl
c 37	11.6	58.0	24	4	US-08-793-937-8	Sequence 8, Appl
c 38	11.6	58.0	24	4	US-09-235-283-8	Sequence 8, Appl
c 39	11.6	58.0	25	2	US-08-859-998-709	Sequence 709, App
c 40	11.6	58.0	26	2	US-08-859-998-496	Sequence 496, App
c 41	11.6	58.0	48	2	US-08-811-492-124	Sequence 124, Appl
c 42	11.4	57.0	18	1	US-08-234-613-43	Sequence 43, Appl
c 43	11.4	57.0	18	1	US-08-484-192-115	Sequence 115, App
c 44	11.4	57.0	18	2	US-08-237-973-56	Sequence 56, Appl
c 45	11.4	57.0	20	3	US-09-418-641-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1
US-09-377-310-14
; Sequence 14, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-14

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.072;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggttagggatggtgcgtca 20
Db 1 ggttagggatggtgcgtca 20

RESULT 2
US-09-377-310-34
; Sequence 34, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-34

Query Match 75.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ttaggagtggtccg 17
Db 1 ttaggagtggtccg 15

RESULT 3

US-09-004-731-58

; Sequence 58, Application US/09004731

; Patent No. 6177258

; GENERAL INFORMATION:

; APPLICANT: Wu Hunter, Shirley

; APPLICANT: Stiegler, Gary

; APPLICANT: Gaines, Patrick J.

; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID

; TITLE OF INVENTION: MOLECULES AND USES THEREOF

; NUMBER OF SEQUENCES: 103

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheridan Ross P.C.

; STREET: 1700 Lincoln Street, Suite 3500

; CITY: Denver

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80203

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/004,731

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/749,699

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Connell, Gary J.

; REGISTRATION NUMBER: 32,020

; REFERENCE/DOCKET NUMBER: 2618-25-C3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (303) 863-9700

; TELEFAX: (303) 863-0223

; INFORMATION FOR SEQ ID NO: 58:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 34 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (primer)

US-09-004-731-58

Query Match 71.0%; Score 14.2; DB 4; Length 34;
Best Local Similarity 84.2%; Pred. No. 71;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gttaggagtggtccgtca 20
Db 13 GTTAGCGATCGTCCCGTCA 31

RESULT 4

US-08-749-699-58

; Sequence 58, Application US/08749699

; Patent No. 6210920
; GENERAL INFORMATION:
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,699
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (primer)
US-08-749-699-58

Query Match 71.0%; Score 14.2; DB 4; Length 34;
Best Local Similarity 84.2%; Pred. No. 71;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gttaggagtggtccgtca 20
Db 13 GTTAGCGATCGTCCCGTCA 31

RESULT 5

US-08-460-344-18/c

; Sequence 18, Application US/08460344

; Patent No. 5654418

; GENERAL INFORMATION:

; APPLICANT: SHEINNESS, Diana K.

; APPLICANT: ADAMS, Trevor H.

; APPLICANT: STAMM, Michael R.

; APPLICANT: CANGELOSI, Gerard A.

; APPLICANT: BRITSCHGL, Theresa B.

; APPLICANT: DIX, Connie K.

; TITLE OF INVENTION: METHODS AND DIAGNOSTIC KITS USEFUL FOR

; TITLE OF INVENTION: DETECTING MICROORGANISMS ASSOCIATED WITH VAGINAL

; TITLE OF INVENTION: INFECTIONS

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Kourie and Crew

; STREET: Steuart Street Tower, One Market Plaza

; CITY: San Francisco

; STATE: California

; COUNTRY: US

; ZIP: 94105-1493

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,344
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/133,598
; FILING DATE: 08-OCT-1993
; APPLICATION NUMBER: US 07/896,094
; FILING DATE: 29-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/600,334
; FILING DATE: 19-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 11652-73-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..24
; OTHER INFORMATION: /standard_name= "GV003"
US-08-460-344-18

Query Match 69.0%; Score 13.8; DB 1; Length 24;
Best Local Similarity 88.2%; Pred. No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 ttaggagtggtgcgcgc 19
|| ||||| |||||
Db 18 TTTGGGATGGAGCCGTC 2

RESULT 6
US-08-133-598A-18/c
; Sequence 18, Application US/08133598A
; Patent No. 5700636
; GENERAL INFORMATION:
; APPLICANT: SHEINESS, Diana K.
; APPLICANT: ADAMS, Trevor H.
; APPLICANT: STAMM, Michael R.
; APPLICANT: CANGELOSI, Gerard A.
; APPLICANT: BRITSCHGI, Theresa B.
; APPLICANT: DIX, Connie K.
; TITLE OF INVENTION: METHODS AND DIAGNOSTIC KITS USEFUL FOR
; TITLE OF INVENTION: DETECTING MICROORGANISMS ASSOCIATED WITH VAGINAL
; TITLE OF INVENTION: INFECTIONS
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,598A
; FILING DATE: 08-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/896,094
; FILING DATE: 29-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/600,334
; FILING DATE: 19-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 11652-73-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..24
; OTHER INFORMATION: /standard_name= "GV003"
US-08-133-598A-18

Query Match 69.0%; Score 13.8; DB 1; Length 24;
Best Local Similarity 88.2%; Pred. No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 ttaggagtggtgcgcgc 19
|| ||||| |||||
Db 18 TTTGGGATGGAGCCGTC 2

RESULT 7
US-08-886-999-18/c
; Sequence 18, Application US/0886999
; Patent No. 5776694
; GENERAL INFORMATION:
; APPLICANT: SHEINESS, Diana K.
; APPLICANT: ADAMS, Trevor H.
; APPLICANT: STAMM, Michael R.
; APPLICANT: CANGELOSI, Gerard A.
; APPLICANT: BRITSCHGI, Theresa B.
; APPLICANT: DIX, Connie K.
; TITLE OF INVENTION: METHODS AND DIAGNOSTIC KITS USEFUL FOR
; TITLE OF INVENTION: DETECTING MICROORGANISMS ASSOCIATED WITH VAGINAL
; TITLE OF INVENTION: INFECTIONS
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/886,999
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/458,319

```
/
/ FILING DATE: 02-JUN-1995
/ APPLICATION NUMBER: US 08/133,598
/ FILING DATE: 08-OCT-1993
/ APPLICATION NUMBER: US 07/896,094
/ FILING DATE: 29-MAY-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/600,334
/ FILING DATE: 19-OCT-1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Garrett-Wackowski, Eugenia
/ REGISTRATION NUMBER: 37,330
/ REFERENCE/DOCKET NUMBER: 11652-73-2
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 543-9600
/ TELEFAX: (415) 543-5043
/ INFORMATION FOR SEQ ID NO: 18:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 24 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1..24
/ OTHER INFORMATION: /standard_name= "GV003"
/
/ US-08-886-999-18
/
/ Query Match 69.0%; Score 13.8; DB 1; Length 24;
/ Best Local Similarity 88.2%; Pred. No. 1.1e+02;
/ Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
/ QY 3 ttaggatggtgcgc 19
/ || ||||| |||||
/ Db 18 TTTGGGATGGAGCCGTC 2
/
/ RESULT 8
/ PCT-US93-05085-18/c
/ APPLICATION: PCT/US9305085
/ GENERAL INFORMATION:
/ APPLICANT: MICROPROBE CORPORATION
/ TITLE OF INVENTION: METHODS AND PHARMACEUTICAL KITS USEFUL
/ FOR DETECTING MICROORGANISMS ASSOCIATED WITH VAGINAL
/ INFECTIONS
/ NUMBER OF SEQUENCES: 57
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Townsend and Townsend Khourie and Crew
/ STREET: Steuart Street Tower, One Market Plaza
/ CITY: San Francisco
/ STATE: California
/ COUNTRY: US
/ ZIP: 94105-1493
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US93/05085
/ FILING DATE: 28-MAY-1993
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Weber, Ellen L.
/ REGISTRATION NUMBER: 32,762
/ REFERENCE/DOCKET NUMBER: 11652-73-1PC
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 543-9600
/ TELEFAX: (415) 543-5043
/ INFORMATION FOR SEQ ID NO: 18:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 24 base pairs
```

```
/
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ PCT-US93-05085-18
/
/ Query Match 69.0%; Score 13.8; DB 5; Length 24;
/ Best Local Similarity 88.2%; Pred. No. 1.1e+02;
/ Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
/ QY 3 ttaggatggtgcgc 19
/ || ||||| |||||
/ Db 18 TTTGGGATGGAGCCGTC 2
/
/ RESULT 9
/ US-08-219-012-36
/ Sequence 36, Application US/08219012
/ Patent No. 5543293
/ GENERAL INFORMATION:
/ APPLICANT: Larry Gold
/ APPLICANT: Diane Tasset
/ TITLE OF INVENTION: Ligands of Thrombin
/ NUMBER OF SEQUENCES: 92
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Beaton & Swanson, P.C.
/ STREET: 4582 South Ulster Street Parkway, Suite #
/ CITY: Denver
/ STATE: Colorado
/ COUNTRY: USA
/ ZIP: 80237
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
/ COMPUTER: IBM compatible
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WordPerfect 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/219,012
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA: none
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Barry J. Swanson
/ REGISTRATION NUMBER: 33,215
/ REFERENCE/DOCKET NUMBER:
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (303) 850-9900
/ TELEFAX: (303) 850-9401
/ INFORMATION FOR SEQ ID NO: 36:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 30 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-219-012-36
/
/ Query Match 69.0%; Score 13.8; DB 1; Length 30;
/ Best Local Similarity 88.2%; Pred. No. 1.1e+02;
/ Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
/
/ QY 1 ggttagggatggtgcgc 17
/ ||||| ||||| ||
/ Db 11 GGTtagggatggtgcgc 27
/
/ RESULT 10
/ US-08-687-421-224
/ Sequence 224, Application US/08687421
/ Patent No. 6177557
```


GENERAL INFORMATION:
; APPLICANT: Gold, Larry
; APPLICANT: Janjic, Nebojsa
; APPLICANT: Tasset, Diane
; TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF BASIC
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR AND
; TITLE OF INVENTION: THROMBIN
; NUMBER OF SEQUENCES: 445
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,421
; FILING DATE: 08-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,005
; FILING DATE: 10-FEBRUARY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 22-APRIL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/219,012
; FILING DATE: 28-MARCH-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,333
; FILING DATE: 11-NOVEMBER-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX07/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 224:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-687-421-224

Query Match 69.0%; Score 13.8; DB 4; Length 30;
Best Local Similarity 88.2%; Pred. No. 1.1e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gttaggatggtgccg 17
DB 11 GGTAGGATGAGTACCG 27

RESULT 11
US-09-156-807-11/c
; Sequence 11, Application US/09156807
; Patent No. 6030786
; GENERAL INFORMATION:
; APPLICANT: Cowser, Lex M.

; TITLE OF INVENTION: ANTISENSE MODULATION OF RHOX EXPRESSION
; FILE REFERENCE: RTS-0014
; CURRENT APPLICATION NUMBER: US/09/156,807
; CURRENT FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 11
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-156-807-11

Query Match 67.0%; Score 13.4; DB 3; Length 18;
Best Local Similarity 93.3%; Pred. No. 1.7e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gttaggatggtgcc 16
DB 15 GTTGGGATGAGTGC 1

RESULT 12
US-09-156-807-12/c
; Sequence 12, Application US/09156807
; Patent No. 6030786
; GENERAL INFORMATION:
; APPLICANT: Cowser, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF RHOX EXPRESSION
; FILE REFERENCE: RTS-0014
; CURRENT APPLICATION NUMBER: US/09/156,807
; CURRENT FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 12
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-156-807-12

Query Match 67.0%; Score 13.4; DB 3; Length 18;
Best Local Similarity 93.3%; Pred. No. 1.7e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 gttaggatggtgcc 16
DB 16 GTTGGGATGAGTGC 2

RESULT 13
US-08-961-083-339/c
; Sequence 339, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 339:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-083-339

Query Match 63.0%; Score 12.6; DB 3; Length 37;
Best Local Similarity 78.9%; Pred. No. 4.7e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 gtagggatggtgccgtca 20
| ||| ||||| |||||
Db 19 GCTAGCGATGGATCCGTCA 1

RESULT 14
5240845-45/c
; Patent No. 5240845
; APPLICANT: FUJII, SEISUO; TAKADA, KAORUKO; KATANO, TAMIKI;
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,
; TSUTOMU
; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS
; NUMBER OF SEQUENCES: 65
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/549,049
; FILING DATE: 06-JUL-1990
; SEQ ID NO: 45:
; LENGTH: 50
5240845-45

Query Match 63.0%; Score 12.6; DB 6; Length 50;
Best Local Similarity 78.9%; Pred. No. 4.9e+02;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 gtagggatggtgccgtca 20
| ||| ||||| |||||
Db 47 GATACGGTTGGTGTCTGTCA 29

RESULT 15
US-09-418-641-75
; Sequence 75, Application US/09418641A
; Patent No. 6124133
; GENERAL INFORMATION:
; APPLICANT: Jennifer K. Taylor
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF FRA-1 EXPRESSION
; FILE REFERENCE: RTS-0105
; CURRENT APPLICATION NUMBER: US/09/418,641A
; CURRENT FILING DATE: 1999-10-15
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 75
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Antisense Oligonucleotide
US-09-418-641-75

Query Match 61.0%; Score 12.2; DB 3; Length 20;
Best Local Similarity 82.4%; Pred. No. 6.9e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 tagggatggtgccgtca 20
| |||| ||||| |||||
Db 1 tagggaggggtgtgggtca 17

Search completed: October 2, 2001, 16:03:50
Job time: 14594 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:48 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-15
Perfect score: 20
Sequence: 1 tgttggtttccaatcgacc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pl1:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_om:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vil:*
- 59: gb_vil2:*
- 60: gb_vtg1:*
- 61: gb_vtg2:*
- 62: gb_vtg3:*
- 63: gb_vtg4:*
- 64: gb_vtg5:*
- 65: gb_vtg6:*
- 66: gb_vtg7:*
- 67: gb_vtg8:*
- 68: gb_vtg9:*
- 69: gb_vtg10:*
- 70: gb_vtg11:*
- 71: gb_vtg12:*
- 72: gb_vtg13:*
- 73: gb_vtg14:*
- 74: gb_vtg15:*
- 75: gb_vtg16:*
- 76: gb_vtg17:*
- 77: gb_vtg18:*
- 78: gb_vtg19:*
- 79: gb_vtg20:*
- 80: gb_vtg21:*
- 81: gb_vtg22:*
- 82: gb_vtg23:*
- 83: gb_vtg24:*
- 84: gb_vtg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_ro1:*
- 95: gb_ro2:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB ID	Description
		Match	Length	%		
C 1	13.6	68.0	36	10	I09246	I09246 Sequence 12
C 2	13.6	68.0	37	10	I09244	I09244 Sequence 10
C 3	13.6	68.0	38	10	I09247	I09247 Sequence 13
C 4	13.4	67.0	30	9	AR109705	AR109705 Sequence
C 5	13.2	66.0	26	9	A56687	A56687 Sequence 6
C 6	13	65.0	19	10	E05267	E05267 Primer for
C 7	12.4	62.0	21	9	AX023416	AX023416 Sequence
C 8	12.4	62.0	24	9	AX052544	AX052544 Sequence

c 9 12 60.0 26 10 E32827 Novel compo
 c 10 12 60.0 30 10 I69336 Sequence 8
 c 11 12 60.0 42 9 AX019232 Sequence
 c 12 12 60.0 48 10 I08124 Sequence 6
 c 13 11.8 59.0 24 9 A43528 Sequence 8
 c 14 11.8 59.0 24 9 AR052465 Sequence
 c 15 11.8 59.0 35 9 AR011952 Sequence
 c 16 11.8 59.0 35 9 AR093800 Sequence
 c 17 11.6 58.0 27 9 AX023890 Sequence
 c 18 11.6 58.0 27 9 S59402 TCR beta -T
 c 19 11.6 58.0 30 9 A21431
 c 20 11.6 58.0 30 9 A43129
 c 21 11.6 58.0 34 9 AR091842 Sequence
 c 22 11.6 58.0 41 9 AR109138
 c 23 11.6 58.0 50 10 E25881 Blastocyst
 c 24 11.4 57.0 18 9 A26514 PT-26I muta
 c 25 11.4 57.0 21 10 AX097321 Sequence
 c 26 11.4 57.0 28 9 AR090859 Sequence
 c 27 11.4 57.0 29 10 AX099592 Sequence
 c 28 11.2 56.0 21 10 AX082611 Sequence
 c 29 11.2 56.0 21 10 AX082613 Sequence
 c 30 11.2 56.0 21 10 AX082616 Sequence
 c 31 11.2 56.0 21 10 AX082618 Sequence
 c 32 11.2 56.0 25 9 AR094119 Sequence
 c 33 11.2 56.0 26 9 AR038156 Sequence
 c 34 11.2 56.0 28 9 A14501 Group 8 Hyb
 c 35 11.2 56.0 28 9 A14502 Group 8 bis
 c 36 11.2 56.0 28 9 A14503 Group 8 ter
 c 37 11.2 56.0 28 9 A14504 Group 8 qua
 c 38 11.2 56.0 28 9 A14559 Hybridizati
 c 39 11.2 56.0 30 9 A21427
 c 40 11.2 56.0 32 10 AX085274
 c 41 11.2 56.0 34 9 AR012356 Sequence
 c 42 11.2 56.0 34 9 AX006548 Sequence
 c 43 11.2 56.0 34 9 AX006652 Sequence
 c 44 11.2 56.0 34 9 AX030256 Sequence
 c 45 11.2 56.0 34 9 AX049934 Sequence

ALIGNMENTS

RESULT 1
 I09246/c I09246 Sequence 36 bp PAT 02-DEC-1994
 LOCUS Sequence 12 from Patent WO 8901940.
 DEFINITION I09246
 ACCESSION I09246
 VERSION I09246.1 GI:588030
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 36)
 AUTHORS Fisher, R.A., Gilbert, W., Sato, V.L., Flavell, R.A., Maraganore, J.M.
 TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
 PRODUCING SOLUBLE T4 PROTEINS
 JOURNAL Patent: WO 8901940-A 12 09-MAR-1989;
 FEATURES Location/Qualifiers
 source 1..36
 BASE COUNT 9 a 9 c 10 g 8 t
 ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 36;
 Best Local Similarity 80.0%; Pred. No. 8.1e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 tgttggtttccaatcgacc 20
 ||||| ||||| |||||
 Db 28 TGTGGATTCCAGCAGGACC 9

RESULT 2
 I09244/c I09244 Sequence 37 bp PAT 02-DEC-1994
 LOCUS Sequence 10 from Patent WO 8901940.
 DEFINITION I09244
 ACCESSION I09244
 VERSION I09244.1 GI:588028
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 37)
 AUTHORS Fisher, R.A., Gilbert, W., Sato, V.L., Flavell, R.A., Maraganore, J.M.
 TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
 PRODUCING SOLUBLE T4 PROTEINS
 JOURNAL Patent: WO 8901940-A 10 09-MAR-1989;
 FEATURES Location/Qualifiers
 source 1..37
 BASE COUNT 11 a 9 c 10 g 7 t
 ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 37;
 Best Local Similarity 80.0%; Pred. No. 8.1e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 tgttggtttccaatcgacc 20
 ||||| ||||| |||||
 Db 28 TGTGGATTCCAGCAGGACC 9

RESULT 3
 I09247 I09247 Sequence 38 bp PAT 02-DEC-1994
 LOCUS Sequence 13 from Patent WO 8901940.
 DEFINITION I09247
 ACCESSION I09247
 VERSION I09247.1 GI:588031
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 38)
 AUTHORS Fisher, R.A., Gilbert, W., Sato, V.L., Flavell, R.A., Maraganore, J.M.
 TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
 PRODUCING SOLUBLE T4 PROTEINS
 JOURNAL Patent: WO 8901940-A 13 09-MAR-1989;
 FEATURES Location/Qualifiers
 source 1..38
 BASE COUNT 8 a 9 c 12 g 9 t
 ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 38;
 Best Local Similarity 80.0%; Pred. No. 8.1e+03;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 tgttggtttccaatcgacc 20
 ||||| ||||| |||||
 Db 15 TGTGGATTCCAGCAGGACC 34

RESULT 4
 AR109705 AR109705 Sequence 30 bp DNA PAT 14-FEB-2001
 LOCUS Sequence 129 from patent US 6114139.
 DEFINITION AR109705
 ACCESSION AR109705
 VERSION AR109705.1 GI:12825981
 KEYWORDS
 SOURCE Unknown.

OS	Artificial gene
OC	Artificial sequence; Genes.
PN	JP 1993192200-A/37
PD	03-AUG-1993
Pf	19-AUG-1991 JP 1991230839
PR	20-AUG-1990 JP 90P 217067
PI	OKAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI FUKUSHIMA MICHIO,
PC	C1201/70,C1201/68;
CC	strandedness: Single;
CC	topology: Linear;
CC	hypothetical: No;
CC	anti-sense: No;
FEATURES	Location/Qualifiers
source	1..19
/organism=	"synthetic construct"
/db_xref=	"taxon:32630" 3 t
BASE COUNT	7 a 4 c 5 g
ORIGIN	
Query Match	65.0%; Score 13; DB 10; Length 19;
Best Local Similarity	100.0%; Pred. No. 1.8e+04;
Matches 13; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 tgttggtttccaa 13
Dd	
Bd	14 TGTGGTTTCCA 2
RESULT	7
LOCUS	AX023416 21 bp DNA PAT 15-SEP-2000
DEFINITION	Sequence 31 from Patent WO0014217.
ACCESSION	AX023416
VERSION	AX023416.1 GI:10183816
KEYWORDS	synthetic construct.
SOURCE	artificial construct.
ORGANISM	artificial sequence.
REFERENCE	1 (bases 1 to 21)
AUTHORS	Lipford,G.B., Heeg,K. and Wagner,H.
TITLE	G-motif oligonucleotides and uses thereof
JOURNAL	Patent: WO 0014217-A 31 16-MAR-2000;
	LIPFORD GRAYSON B (DE) ; HEEG KLAUS (DE) ; WAGNER HERMANN (DE) ; CPG IMMUNOPHARMACEUTICALS GMBH (DE)
FEATURES	Location/Qualifiers
source	1..21
/organism=	"synthetic construct"
/db_xref=	"taxon:32630"
/note=	"synthetic, no natural origin"
BASE COUNT	2 a 5 c 3 g 11 t
ORIGIN	
Query Match	62.0%; Score 12.4; DB 9; Length 21;
Best Local Similarity	92.9%; Pred. No. 4.1e+04;
Matches 13; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
Qy	1 tgttggtttccaat 14
Dd	
Bd	8 TGTGGTTTCCTAT 21
RESULT	8
LOCUS	AX052544 24 bp DNA PAT 12-JAN-2001
DEFINITION	Sequence 6 from Patent WO0071668.
ACCESSION	AX052544
VERSION	AX052544.1 GI:12226774
KEYWORDS	synthetic construct.
SOURCE	

```

ORGANISM      synthetic construct
REFERENCE      artificial sequence.
AUTHORS        1 (bases 1 to 24)
TITLE          Rhoades,D.M.
JOURNAL        Method for identifying components involved in signal transduction
                pathways in higher plants
                Patent: WO 0071668-A 6 30-NOV-2000;
                The Board of Regents of the University of Nebraska (US)
FEATURES       Location/Qualifiers
source         1..24
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="NCOI Primer"
BASE COUNT     6 a      5 c      5 g      8 t
ORIGIN
Query Match    62.0%; Score 12.4; DB 9; Length 24;
Best Local Similarity 92.9%; Pred. No. 4.1e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 tgggttccaatcg 17
||||| |||||
Db 11 TGGTTCAAATCG 24

RESULT 9
E32827/c      26 bp      DNA      PAT      07-FEB-2001
LOCUS         Novel compound.
DEFINITION    E32827
ACCESSION     E32827.1 GI:13026860
VERSION       JP 1999137263-A/3.
KEYWORDS      unidentified.
SOURCE        unidentified.
ORGANISM      unclassified.
REFERENCE      1 (bases 1 to 26)
AUTHORS        Nicola,G.W.
TITLE          Novel compound
JOURNAL        Patent: JP 1999137263-A 3 25-MAY-1999;
                SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM CORP PUBLIC LTD CO
COMMENT        OS Unidentified
                PN JP 1999137263-A/3
                PD 25-MAY-1999
                PR 18-JUN-1998 JP 1998210193
                PF 20-JUN-1997 US 08/880101
                PI NICOLA GAIL WALLIS
                PC C12N15/09,A61K35/74,A61K35/76,A61K39/09,A61K45/00,A61K48/00,
                PC C07K14/315,
                PC C07K16/12,C12Q1/68,G01N33/53,G01N33/566//A61K39/04,(C12N15/09,
                PC C12R1:46),
                PC C12N15/00,(C12N15/00,C12R1:46)
                CC Strandedness: Single;
                CC Topology: Linear;
                FH Key      Location/Qualifiers
                FT source 1..26
                FT         /organism="Unidentified".
                FEATURES
                source    1..26
                Location/Qualifiers
                /organism="unidentified"
                /db_xref="taxon:32644"
BASE COUNT     8 a      7 c      1 g      10 t
ORIGIN

Query Match    60.0%; Score 12; DB 10; Length 26;
Best Local Similarity 75.0%; Pred. No. 6.9e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tgggtttccaatcgacc 20
||||| |||||
Db 25 TGGTGTATATAATTGGAGC 6

ORGANISM      synthetic construct
REFERENCE      artificial sequence.
AUTHORS        1 (bases 1 to 24)
TITLE          Rhoades,D.M.
JOURNAL        Method for identifying components involved in signal transduction
                pathways in higher plants
                Patent: WO 0071668-A 6 30-NOV-2000;
                The Board of Regents of the University of Nebraska (US)
FEATURES       Location/Qualifiers
source         1..24
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="NCOI Primer"
BASE COUNT     6 a      5 c      5 g      8 t
ORIGIN
Query Match    62.0%; Score 12.4; DB 9; Length 24;
Best Local Similarity 92.9%; Pred. No. 4.1e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 tgggttccaatcg 17
||||| |||||
Db 11 TGGTTCAAATCG 24

RESULT 9
E32827/c      26 bp      DNA      PAT      07-FEB-2001
LOCUS         Novel compound.
DEFINITION    E32827
ACCESSION     E32827.1 GI:13026860
VERSION       JP 1999137263-A/3.
KEYWORDS      unidentified.
SOURCE        unidentified.
ORGANISM      unclassified.
REFERENCE      1 (bases 1 to 26)
AUTHORS        Nicola,G.W.
TITLE          Novel compound
JOURNAL        Patent: JP 1999137263-A 3 25-MAY-1999;
                SMITHKLINE BEECHAM CORP,SMITHKLINE BEECHAM CORP PUBLIC LTD CO
COMMENT        OS Unidentified
                PN JP 1999137263-A/3
                PD 25-MAY-1999
                PR 18-JUN-1998 JP 1998210193
                PF 20-JUN-1997 US 08/880101
                PI NICOLA GAIL WALLIS
                PC C12N15/09,A61K35/74,A61K35/76,A61K39/09,A61K45/00,A61K48/00,
                PC C07K14/315,
                PC C07K16/12,C12Q1/68,G01N33/53,G01N33/566//A61K39/04,(C12N15/09,
                PC C12R1:46),
                PC C12N15/00,(C12N15/00,C12R1:46)
                CC Strandedness: Single;
                CC Topology: Linear;
                FH Key      Location/Qualifiers
                FT source 1..26
                FT         /organism="Unidentified".
                FEATURES
                source    1..26
                Location/Qualifiers
                /organism="unidentified"
                /db_xref="taxon:32644"
BASE COUNT     8 a      7 c      1 g      10 t
ORIGIN

```

```

RESULT 10
I69336/c      30 bp      DNA      PAT      04-FEB-1998
LOCUS         Sequence 8 from patent US 5677172.
DEFINITION    I69336
ACCESSION     I69336
VERSION       I69336.1 GI:2831458
KEYWORDS      Unknown.
SOURCE        Unknown.
ORGANISM      Unclassified.
REFERENCE      1 (bases 1 to 30)
AUTHORS        Makarow,M.
TITLE          Method for production of proteins in yeast
JOURNAL        Patent: US 5677172-A 8 14-OCT-1997;
FEATURES       Location/Qualifiers
source         1..30
                /organism="unknown"
BASE COUNT     6 a      9 c      8 g      7 t
ORIGIN
Query Match    60.0%; Score 12; DB 10; Length 30;
Best Local Similarity 75.0%; Pred. No. 6.9e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tgggtttccaatcgacc 20
||||| |||||
Db 24 TGGTGTTCACAAAGGTACC 5

RESULT 11
AX019232/c    42 bp      DNA      PAT      07-SEP-2000
LOCUS         Sequence 12 from Patent WO9941390.
DEFINITION    AX019232
ACCESSION     AX019232
VERSION       AX019232.1 GI:10043261
KEYWORDS      synthetic construct.
SOURCE        synthetic construct.
ORGANISM      artificial sequence.
REFERENCE      1 (bases 1 to 42)
AUTHORS        Huels,C., Muellner,S., Boehnisch,B. and Gallert,K.C.
TITLE          Expression vector for the production of dead proteins
JOURNAL        Patent: WO 9941390-A 12 19-AUG-1999;
                HUELS CHRISTOPH (DE); MUELLNER STEFAN (DE); AVENTIS RES & TECH GMBH
                & CO (DE); BOEHNISCH BRITTA (DE); GALLERT KARL CHRISTIAN (DE)
FEATURES       Location/Qualifiers
source         1..42
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="p135-NT5C"
BASE COUNT     12 a     13 c      6 g     11 t
ORIGIN
Query Match    60.0%; Score 12; DB 9; Length 42;
Best Local Similarity 75.0%; Pred. No. 6.8e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tgggtttccaatcgacc 20
||||| |||||
Db 32 TGATGATGTCCTCGGATC 13

RESULT 12
I08124/c      48 bp      PAT      02-DEC-1994
LOCUS         Sequence 6 from Patent EP 0334004.
DEFINITION    I08124
ACCESSION     I08124
VERSION       I08124.1 GI:589165
KEYWORDS      Unknown.
SOURCE        Unknown.

```



```
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 48)
AUTHORS Smith,J.A. and Lee,F-J.S.
TITLE Isolation, purification, characterization, cloning and sequencing
of N alpha-acetyltransferase
JOURNAL Patent: EP 0334004-A1 6 27-SEP-1989;
FEATURES Location/Qualifiers
source
1..48
/organism="unknown"
BASE COUNT 14 a 10 c 13 g 11 t
ORIGIN

Query Match 60.0%; Score 12; DB 10; Length 48;
Best Local Similarity 75.0%; Pred. No. 6.8e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 tgggtttccaatcgacc 20
Db 43 TGCTGCTTACCACATCTGACC 24

RESULT 13
LOCUS A43528 24 bp DNA PAT 06-MAR-1997
DEFINITION Sequence 8 from Patent WO9507981.
ACCESSION A43528
VERSION A43528.1 GI:2298716
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequence.
REFERENCE 1 (bases 1 to 24)
AUTHORS Schweighoffer,F. and Tocque,B.
TITLE GRB3-3 GENE, VARIANTS AND USES THEREOF
JOURNAL Patent: WO 9507981-A 8 23-MAR-1995;
COMMENT RHONE POULENC RORER SA (FR)
Other publication PL 313445 960708
Other publication NO 960965 960308
Other publication FI 961202 960314
Other publication CA 2169338 950323
Other publication ZA 9407059 950518
Other publication AU 6724794 950403
Other publication FR 2710074 950324.
FEATURES Location/Qualifiers
source
1..24
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 5 a 4 c 6 g 9 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 9; Length 24;
Best Local Similarity 86.7%; Pred. No. 9e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 tgggtttccaatcgga 18
Db 2 TCGTTTCCAAACGGA 16

RESULT 14
LOCUS AR052465 24 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 9 from patent US 5831048.
ACCESSION AR052465
VERSION AR052465.1 GI:5975829
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 24)
```

```
AUTHORS Schweighoffer,F. and Tocque ,B.
TITLE Grb3-3 CDNA and polypeptides
JOURNAL Patent: US 5831048-A 9 03-NOV-1998;
FEATURES Location/Qualifiers
source
1..24
/organism="unknown"
BASE COUNT 5 a 4 c 6 g 9 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 9; Length 24;
Best Local Similarity 86.7%; Pred. No. 9e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 tgggtttccaatcgga 18
Db 2 TCGTTTCCAAACGGA 16

RESULT 15
LOCUS AR011952 35 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 29 from patent US 5763177.
ACCESSION AR011952
VERSION AR011952.1 GI:3969942
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 35)
AUTHORS Gold,L., Willis,M., Koch,T., Ringquist,S., Jensen,K. and
Atkinson,B.
TITLE Systematic evolution of ligands by exponential enrichment:
photoselection of nucleic acid ligands and solution sele
JOURNAL Patent: US 5763177-A 29 09-JUN-1998;
FEATURES Location/Qualifiers
source
1..35
/organism="unknown"
BASE COUNT 8 a 8 c 6 g 13 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 9; Length 35;
Best Local Similarity 86.7%; Pred. No. 9e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tgggtttccaatc 15
Db 18 TCGTTTGTTCACATC 32
```

Search completed: October 2, 2001, 15:56:49
Job time: 14173 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:02 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-15

Perfect score: 20

Sequence: 1 tttgtgttccaatcgacc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
C 1	12.8	64.0	46	12	A8790759	vw18g12.r	AZ790759 vw18g12.r
C 2	12.2	61.0	29	250	A2388877	2M0134E10	AZ388877 2M0134E10
C 3	12.2	61.0	30	249	AZ768472	1M0568C10	AZ768472 1M0568C10
C 4	12	60.0	28	242	AZ380625		AZ380625 1M0136G10
C 5	12	60.0	37	258	TA379B06Q		AL497861 T. brucei
C 6	12	60.0	40	245	AZ537263		AZ537263 AST-2P031
C 7	12	60.0	43	13	AA895268	vx49C02.r	AA895268 vx49C02.r
C 8	12	60.0	43	13	AA895268		AA895268 vx49C02.r
C 9	12	60.0	43	241	AZ319753	1M0039B02	AZ319753 1M0039B02
C 10	11.8	59.0	102	50	AU107355		AU107355 AU107355
C 11	11.6	58.0	19	241	AZ327390	1M0050L09	AZ327390 1M0050L09
C 12	11.4	57.0	27	250	AZ822425	2M0095M09	AZ822425 2M0095M09
C 13	11.4	57.0	46	23	A1663836	uJ06E03.x	A1663836 uJ06E03.x
C 14	11.4	57.0	50	117	AW537238	G0114A01	AW537238 G0114A01
C 15	11.2	56.0	40	10	AA652673	ns69F11.s	AA652673 ns69F11.s
C 16	11	55.0	43	250	AZ822833	2M0096J23	AZ822833 2M0096J23
C 17	11	55.0	36	243	AA043719	zk61f12.s	AA043719 zk61f12.s
C 18	11	55.0	39	249	AZ427281	1M0209H10	AZ427281 1M0209H10
C 19	11	55.0	40	24	AZ763518	1M0557C02	AZ763518 1M0557C02
C 20	11	55.0	43	189	TA783759	tu45d08.x	TA783759 tu45d08.x
C 21	11	55.0	50	24	T63717	yc16f09.r1	T63717 yc16f09.r1
C 22	11	55.0	40	24	AI750109	at27e02.x	AI750109 at27e02.x
C 23	11	55.0	50	107	AU103004		AU103004 AU103004
C 24	10.8	54.0	22	246	AU103668		AU103668 AU103668
C 25	10.8	54.0	22	246	AZ623308	1M0460F21	AZ623308 1M0460F21
C 26	10.8	54.0	37	249	AZ803825	2M0064I24	AZ803825 2M0064I24
C 27	10.8	54.0	40	166	AA079070	zm95a02.s	AA079070 zm95a02.s
C 28	10.8	54.0	44	244	BE377955	601230I33	BE377955 601230I33
C 29	10.8	54.0	44	244	AZ479665	1M0300D01	AZ479665 1M0300D01
C 30	10.8	54.0	46	1	AZ791420	2M0041Q14	AZ791420 2M0041Q14
C 31	10.8	54.0	46	155	AA037870	zf04d10.r	AA037870 zf04d10.r
C 32	10.8	54.0	46	155	C02388	HUNG0001088	C02388 HUNG0001088
C 33	10.6	53.0	22	246	AZ601191	ym10f05.r1	AZ601191 ym10f05.r1
C 34	10.6	53.0	22	249	AZ787023	1M0419C20	AZ787023 1M0419C20
C 35	10.6	53.0	29	249	AZ779056	2M0032B23	AZ779056 2M0032B23
C 36	10.6	53.0	37	20	AI446151	tJ07e11.x	AI446151 tJ07e11.x
C 37	10.6	53.0	42	249	AZ779815	2M0016C08	AZ779815 2M0016C08
C 38	10.6	53.0	45	244	AZ461904	1M0269A01	AZ461904 1M0269A01
C 39	10.6	53.0	45	250	AZ819815	2M0091I08	AZ819815 2M0091I08
C 40	10.6	53.0	47	258	TA220G01P		TA220G01P
C 41	10.6	53.0	49	2	AA115149	T. brucei	AA115149 T. brucei
C 42	10.6	53.0	50	107	AU108011	zl07D09.r	AA115149 zl07D09.r
C 43	10.6	53.0	50	249	AZ794281	2M0048C04	AZ794281 2M0048C04
C 44	10.4	52.0	23	242	AZ350054	1M0087L18	AZ350054 1M0087L18
C 45	10.4	52.0	28	241	AZ324675	1M0046A14	AZ324675 1M0046A14

ALIGNMENTS

RESULT	1
AA790759/c	
LOCUS	
DEFINITION	
IMAGE:	1244230 5' similar to TR:O02668 INTER-ALPHA-INHIBITOR HEAVY-CHAIN H2. ; mRNA sequence.
ACCSSION	
VERSION	
KEYWORDS	
ORGANISM	
SOURCE	
house mouse.	
Mus musculus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 46)	
REFERENCE	
AUTHORS	
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,	
EST	
GI:2850879	
mRNA	
Soares,mammary.gland.NBMMG Mus musculus cDNA clone	
O02668	
06-FEB-1998	

TITLE	JOURNAL	COMMENT
1. The Role of the State in the Development of the Economy	Journal of Economic Surveys	See page 100
2. The Impact of Globalization on the Labor Market	Journal of International Economics	See page 105
3. The Effect of Trade Liberalization on the Environment	Journal of Environmental Economics	See page 110
4. The Role of the State in the Development of the Environment	Journal of Environmental Economics	See page 115
5. The Impact of Globalization on the Environment	Journal of International Economics	See page 120
6. The Effect of Trade Liberalization on the Environment	Journal of Environmental Economics	See page 125
7. The Role of the State in the Development of the Environment	Journal of Environmental Economics	See page 130
8. The Impact of Globalization on the Environment	Journal of International Economics	See page 135
9. The Effect of Trade Liberalization on the Environment	Journal of Environmental Economics	See page 140
10. The Role of the State in the Development of the Environment	Journal of Environmental Economics	See page 145

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HHMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact
 IMAGE Consortium (info@image.llnl.gov) for further information
 MGI:657918

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.

FEATURES
SOURCE

```

1. 46
2. 13 t
3. 9 q
4. 16 c
5. 8 a
6. 13 t
7. 9 q
8. 16 c
9. 8 a
10. 13 t
11. 9 q
12. 16 c
13. 8 a
14. 13 t
15. 9 q
16. 16 c
17. 8 a
18. 13 t
19. 9 q
20. 16 c
21. 8 a
22. 13 t
23. 9 q
24. 16 c
25. 8 a
26. 13 t
27. 9 q
28. 16 c
29. 8 a
30. 13 t
31. 9 q
32. 16 c
33. 8 a
34. 13 t
35. 9 q
36. 16 c
37. 8 a
38. 13 t
39. 9 q
40. 16 c
41. 8 a
42. 13 t
43. 9 q
44. 16 c
45. 8 a
46. 13 t
47. 9 q
48. 16 c
49. 8 a
50. 13 t
51. 9 q
52. 16 c
53. 8 a
54. 13 t
55. 9 q
56. 16 c
57. 8 a
58. 13 t
59. 9 q
60. 16 c
61. 8 a
62. 13 t
63. 9 q
64. 16 c
65. 8 a
66. 13 t
67. 9 q
68. 16 c
69. 8 a
70. 13 t
71. 9 q
72. 16 c
73. 8 a
74. 13 t
75. 9 q
76. 16 c
77. 8 a
78. 13 t
79. 9 q
80. 16 c
81. 8 a
82. 13 t
83. 9 q
84. 16 c
85. 8 a
86. 13 t
87. 9 q
88. 16 c
89. 8 a
90. 13 t
91. 9 q
92. 16 c
93. 8 a
94. 13 t
95. 9 q
96. 16 c
97. 8 a
98. 13 t
99. 9 q
100. 16 c

```

BASE COUNT	8 a	16 c	9 g	13 t
ORIGIN				

Query Match	64.0%;	Score 12.8;	DB 12;	Length 46;
Best Local Similarity	87.5%;	Pred. No. 1.2e+04;		
Matches 14;	Conservative	0;	Mismatches 2;	Indels 0

Qy 2 gttgggttccaatcgg 17
 || ||||||| |||||
Db 39 GTGGGTTTCCCATCGG 24

RESULT 2
A2838877/C

LOCUS	AZ838877	29 bp	DNA	GSS	20-FEB-2001
DEFINITION	2M0134E10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0134E10 R, DNA sequence.				

ACCESSION AZ838877
VERSION AZ838877.1 GI:13008701
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

ORGANISM
Eukaryotes: Mus musculus

REFERENCE
1 (bases 1 to 29)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.

AUTHORS

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL
Unpublished (2000)

COMMENT

Contact: Robert B. Weiss
University of Utah
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0134 row: E column: 10
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 29.

FEATURES

source

1. .29
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="UUCG2M0134E10"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g1147321141gb1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

6 a 7 c 5 g 11 t

Query Match 61.0%; Score 12.2; DB 250; Length 29;
Best Local Similarity 82.4%; Pred. No. 2 4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ttgggtttccaatcgac 19
||| || ||| ||| ||| |||

Db 29 TTGATTACCACCGGAC 13

RESULT 3

AZ768472 30 bp DNA GSS 16-FEB-2001
LOCUS 1M0568C10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0568C10 R, DNA sequence.
ACCESSION AZ768472
VERSION AZ768472.1 GI:12887815
KEYWORDS GSS.
SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 30)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL
COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0568 row: C column: 10
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 30.

FEATURES

source

1. .30
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="UUGC1M0568C10"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g1147321141gb1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 7 a 3 c 8 g 12 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 249; Length 30;
Best Local Similarity 82.4%; Pred. No. 2 4e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gttggtttccaatcgga 18

||||| ||| ||| ||| |||

Db 3 GTTGTTTACATTCGA 19

RESULT 4

AZ380625/c 28 bp DNA GSS 02-OCT-2000
LOCUS 1M0136G10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0136G10 R, DNA sequence.

ACCESSION AZ380625
VERSION AZ380625.1 GI:10494325
KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 28)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL
COMMENT

plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0136 row: G column: 10
Seq primer: CACACGAGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 28.

FEATURES
source

1. .28
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0136G10"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gii14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

9 a 3 c 9 g 7 t
Query Match 60.0%; Score 12; DB 242; Length 28;
Best Local Similarity 75.0%; Pred. No. 3e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 tggtaggtttccaatcgacc 20
||||| ||||| ||||| |||||

Db 24 TGTAGCTTCCCATCTGAAC 5

RESULT
TA379B06Q/C

LOCUS TA379B06Q 37 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 379b06, reverse sequence, genomic survey sequence.
ACCESSION AL497861
VERSION AL497861.1 GI:11873583
KEYWORDS GSS.

SOURCE
Trypanosoma brucei.

Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.

REFERENCE
AUTHORS

1 (bases 1 to 37)
Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.
Direct Submission

TITLE

JOURNAL

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nh@sanger.ac.uk
Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org

Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES
source

1. .37
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="379b06"

BASE COUNT 12 a 10 c 6 g 9 t
ORIGIN

Query Match 60.0%; Score 12; DB 258; Length 37;
Best Local Similarity 75.0%; Pred. No. 3.2e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 tggtaggtttccaatcgacc 20
||||| ||||| ||||| |||||

Db 21 TGTGATTTCCAGAGAAC 2

RESULT
AZ537263

LOCUS AZ537263 40 bp DNA GSS 06-NOV-2000
DEFINITION ASP-2P03160 Genetrap PC-3 Human Prostatic Carcinoma Library Homo sapiens genomic 5', DNA sequence.
ACCESSION AZ537263
VERSION AZ537263.1 GI:11114226
KEYWORDS GSS.
SOURCE human.

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 40)
Henkel, G., Liyanage, M., Pratt, E., Huang, D., Riley, M., Bernardino, A., Durick, K. and Pollok, B.

REFERENCE
AUTHORS

Exon-trap tags from a PC-3 GenomeScreen(TM) Library
Unpublished (2000)
Contact: Greg Henkel
Gene Expression
Aurora Biosciences Corp.
11010 Torreyana Road, San Diego, CA 92121, USA
Tel: 8584048436
Fax: 8584046719
Email: henkel@aurorabio.com

TITLE
JOURNAL

Comment
Pools of cells were isolated from a GenomeScreen(TM) library. The library of cells was generated by retroviral integration of a gene tagging element consisting of: 1) A promoterless beta-lactamase preceded by a splice acceptor as a reporter for gene expression; 2) A promoter driving neomycin resistance followed by a splice donor to trap downstream exons. 3' RACE from neomycin gene was performed using total RNA from isolated pools. Output was shotgun bacteria. 5' ends of reported sequences were immediately preceded by splice donor from the trapping construct.

FEATURES
source

1. .40
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone_lib="Genetrap PC-3 Human Prostatic Carcinoma
Library"
/tissue_type="Adenocarcinoma"
/cell_type="Epithelial"
/clone_line="PC-3"
/notes="Organ: Prostate; Vector: pAmp-1; 3' RACE of total
RNA from genetrap pools; shotgun clone in pAmp-1 and used
to transform DH5-alpha competent bacteria."
BASE COUNT      12 a      12 c      4 g      12 t
ORIGIN

```

```

Query Match      60.0%; Score 12; DB 245; Length 40;
Best Local Similarity 75.0%; Pred. No. 3.3e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 ttgttggttccaatcgacc 20
||||| ||||| | |||
Db 5 TGTGTGATTTCCACATGAAC 24

```

```

RESULT 7
AA895268/c 43 bp mRNA EST 06-APR-1998
LOCUS
DEFINITION
IMAGE:1278530 5' similar to TR:002754 O02754 C/EBP ALPHA PROTEIN.
; mRNA sequence.
ACCESSION
AA895268
VERSION
AA895268.1 GI:3031687
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 43)

```

```

REFERENCE
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:670330

```

```

Trace considered overall poor quality
Possible reversed clone: Similarity on wrong strand
Seq primer: -28m3 revl Et from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1..43
/organism="Mus musculus"
/strain="C57BL/6 x CBA"
/db_xref="taxon:10090"
/clone="IMAGE:1278530"
/clone_lib="Stratagene mouse lung 937302"
/sex="female"
/tissue_type="lung"
/dev_stage="6-8 month old"
/lab_host="SOUR (kanamycin resistant)"
/notes="Organ: lung; Vector: pBluescript SK-; Site 1: EcoRI
; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. 6-8 month old female lung and 1.5 year old male lung
were source of mRNA. Average insert size: 1.5 kb; Uni-ZAP
XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3'

```

```

FEATURES
source

```

```

adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTTTT 3'
BASE COUNT      12 a      12 c      14 g      5 t
ORIGIN

```

```

Query Match      60.0%; Score 12; DB 13; Length 43;
Best Local Similarity 75.0%; Pred. No. 3.3e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 ttgttggttccaatcgacc 20
||||| ||||| | |||
Db 32 TGTTCGGTTCCACGCTACC 13

```

```

RESULT 8
A2319753 43 bp DNA GSS 29-SEP-2000
LOCUS
DEFINITION
clone UUGC1M0039B02 R, DNA sequence.
ACCESSION
A2319753
VERSION
A2319753.1 GI:10370842
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 43)

```

```

REFERENCE
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0039 row: B column: 02
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 43.
Location/Qualifiers
1..43
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0039B02"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into

```

```

TITLE
JOURNAL
COMMENT

```

```

FEATURES
source

```


chemically-competent *E. coli* XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

11 a 11 c 13 g 8 t

Query Match 60.0%; Score 12; DB 241; Length 43;
Best Local Similarity 75.0%; Pred. No. 3.3e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tttgtgtttccaatcgacc 20
||||| ||||| |||
Db 8 TGTGGTGTGCAATCCAAAC 27

RESULT 9

AU107355 50 bp mRNA EST 05-APR-2001
LOCUS
DEFINITION AU107355 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
CAS07628, mRNA sequence.

ACCESSION AU107355
VERSION AU107355.1 GI:13556876
KEYWORDS EST.
SOURCE human.

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata
H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo
K., Suyama, A. and Sugano, S.

Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries

Unpublished (2001)

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

Location/Qualifiers

1..50

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CAS07628"

/clone_lib="Sugano Homo sapiens cDNA library"

6 a 15 c 21 g 8 t

BASE COUNT

ORIGIN

Query Match 60.0%; Score 12; DB 107; Length 50;

Best Local Similarity 75.0%; Pred. No. 3.4e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 tttgtgtttccaatcgacc 20
||||| ||||| |||
Db 5 TGATGGTGGCAACGGGCC 24

RESULT 10

AZ327390 19 bp DNA GSS 29-SEP-2000
LOCUS
DEFINITION 1M0050L09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0050L09 R, DNA sequence.

ACCESSION AZ327390
VERSION AZ327390.1 GI:10386092
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS

1 (bases 1 to 19)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmood, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.,
and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0050 row: L column: 09

Seq primer: CACACAGGAACACGCTATGACC

Class: plasmid ends

High quality sequence stop: 19.

Location/Qualifiers

1..19

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0050L09"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/notes="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 (gll4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent *E. coli* XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

2 a 4 c 3 g 10 t

BASE COUNT

ORIGIN

Query Match 59.0%; Score 11.8; DB 241; Length 19;

Best Local Similarity 86.7%; Pred. No. 3.6e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ttgtttccaatcg 17

||||| ||||| |||

Db 5 TTGTGTTCCAATTGG 19

RESULT 11

AZ822425 27 bp DNA GSS 20-FEB-2001

LOCUS

DEFINITION 2M0095M09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC2M0095M09 R, DNA sequence.

ACCESSION AZ822425

VERSION AZ822425.1 GI:12992333

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 27)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.,
 and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0095 row: M column: 09
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 27.

FEATURES

Location/Qualifiers
 1..27
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUC2M0095M09"
 /clone_lib="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 (g114732114[gblAF129072.1]), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 7 a 3 c 8 g 9 t
ORIGIN
 Query Match 58.0%; Score 11.6; DB 250; Length 27;
 Best Local Similarity 77.8%; Pred. No. 5e+04;
 Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1 ttgtgtttcccaatcgga 18
 ||||| ||||| |||||
 Db 1 TGGTGATTTCCTATGGGA 18

RESULT 12
LOCUS AI663836 46 bp mRNA EST 10-MAY-1999
DEFINITION uJ06e03.xl Sugano mouse liver mlla Mus musculus cDNA clone
 IMAGE:1891132 3' similar to SW:ALBU_MERUN O35090 SERUM ALBUMIN
 PRECURSOR. ; mRNA sequence.
ACCESSION AI663836
VERSION AI663836.1 GI:4767419
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 46)
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
 , B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter
 , E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Other_ESTs: uJ06e03.y1

JOURNAL

COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:975456

Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: custom primer used
 High quality sequence stop: 1.

FEATURES

Location/Qualifiers
 1..46
 /organism="Mus musculus"
 /strain="C57BL"
 /db_xref="taxon:10090"
 /clone="IMAGE:1891132"
 /clone_lib="Sugano mouse liver mlla"
 /sex="female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: liver; Vector: pME18S-FL3; Site:1: DraIII
 (CAGCTGTG); Site:2: DraIII (CACCAGTG); 1st strand cDNA
 was primed with an oligo(dT) primer
 [ATGTGGCTTTTGTGCTACTGG]; double-stranded cDNA was
 ligated to a DraIII adaptor [TGTGGCTACTGG], digested
 and cloned into distinct DraIII sites of the pME18S-FL3
 vector (5' site CAGCTGTG, 3' site CACCAGTG). XhoI should
 be used to isolate the cDNA insert. Size selection was
 performed to exclude fragments <1.5kb. Library
 constructed by Dr. Sumio Sugano (University of Tokyo
 Institute of Medical Science). Custom primers for
 sequencing: 5' end primer CTTCGTCTTAAAGCTGCG and 3' end
 primer CGACCTGCAGCTCGAGCACA."

BASE COUNT 13 a 12 c 11 g 10 t
ORIGIN

Query Match 57.0%; Score 11.4; DB 23; Length 46;
 Best Local Similarity 92.3%; Pred. No. 7.1e+04;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ttgtgtttcccaatc 15
 ||||| |||||
 Db 23 TTTGTTTCCATC 11

RESULT

LOCUS AW537238/c 50 bp mRNA EST 31-AUG-2000
DEFINITION G0114A01-3 Mouse E7.5 Embryonic Portion cDNA Library Mus musculus
 cDNA clone G0114A01 3', mRNA sequence.
ACCESSION AW537238
VERSION AW537238.1 GI:7179655
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 50)
 AUTHORS Tanaka,T.S., Jaradat,S.A., Lim,M.K., Kargul,G.J., Wang,X., Grahovac,M.J., Pantano,S., Sano,Y., Piao,Y., Nagaraja,R., Doi,H., Wood,W.H., III, Becker,K.G. and Ko,M.S.H.
 TITLE Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (16), 9127-9132 (2000)
 MEDLINE 20381348
 COMMENT Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@nigun-grc.nia.nih.gov
 Plate: G0114 row: A column: 01
 Seq primer: -21M13 Forward
 High quality sequence stop: 50
 POLYA-Yes.

FEATURES source
 1. 50 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="G0114A01"
 /clone_lib="Mouse E7.5 Embryonic Portion cDNA Library"
 /sex="unknown"
 /dev_stage="7.5dpc Embryo"
 /lab_host="DH10B"
 /note="Vector: pSPORT1 (Gibco/BRL Life Technology); Site_1: SalI; Site_2: NotI; Total RNAs were extracted from 6 Embryo. The double-stranded cDNA was synthesized by Gibco's kit with an oligo(dT) primer [NotI primer-adaptor from Gibco/BRL] [5'-pGACTAGTCTAGATCGCGCGCGCTTTT-3'] from 0.5µg of mRNA. The double-stranded cDNAs were treated with T4 DNA polymerase and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal3 (include SalI sequence). The cDNAs were purified by phenol/chloroform and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Takara's Ex Taq polymerase. Then, the cDNAs were purified by phenol/chloroform and by Centricon 100. The cDNAs were digested with SalI and NotI enzymes. Then, the cDNAs were size selected by Gibco's Size Fractionation Column. The cDNAs were cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by chemical method. The library was constructed by Xiaohong Wang and Minoru S. H. Ko."

BASE COUNT 17 a 14 c 6 g 13 t
 ORIGIN
 Query Match 57.0%; Score 11.4; DB 117; Length 50;
 Best Local Similarity 92.3%; Pred. No. 7.3e+04;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ttggtttccaatc 15
 ||| |||||
 Db 20 TTGGTTCCAATC 8

RESULT 14
 AA652673/c 40 bp mRNA EST 25-NOV-1997
 LOCUS ns69f11.s1 NCI_CGAP_Pr2 Homo sapiens cDNA clone IMAGE:118909
 DEFINITION similar to TR:G1712 G1712 RETROVIRUS RELATED REVERSE TRANSCRIPTASE
 ;contains L1.t1 L1 repetitive element ;, mRNA sequence.
 AA652673
 ACCESSION AA652673.1 GI:2584325
 VERSION
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 40)
 AUTHORS NCI-CGAP
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai, M.D., Michael Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Insert Length: 1233 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 40
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1188909"
 /clone_lib="NCI_CGAP_Pr2"
 /sex="Male"
 /dev_stage="45 years old"
 /lab_host="DH10B"
 /note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krimman."

BASE COUNT 21 a 3 c 9 g 7 t
 ORIGIN
 Query Match 56.0%; Score 11.2; DB 10; Length 40;
 Best Local Similarity 81.2%; Pred. No. 8.9e+04;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ttggtttccaatcgga 18
 ||| |||||
 Db 40 TTCCTTCCAATCNGA 25

RESULT 15
 AZ822833 43 bp DNA GSS 20-FEB-2001
 LOCUS 2M0096J23F Mouse 10kb plasmid UGCCIM library Mus musculus genomic
 DEFINITION clone UGCC2M0096J23 F, DNA sequence.
 AZ822833
 ACCESSION
 VERSION
 KEYWORDS GSS.
 SOURCE house musculus.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 43)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A. and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0096 row: J column: 23
Seq primer: CTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 43.

FEATURES

source

1..43
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0096J23"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42hv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g147321141gb1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

11 a 8 c 8 g 16 t

Query Match 56.0%; Score 11.2; DB 250; Length 43;
Best Local Similarity 81.2%; Pred. NO. 9e+04;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 ttgggttccaatcgga 18
| | | | | | | | | |
Db 21 TGGGTTTACACTCGGA 36

Search completed: October 2, 2001, 15:01:04
Job time: 10833 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:49 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-16
Perfect score: 20
Sequence: 1 ctaggaggagctcagtggtgg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl:*
- 1: gb_ba1:*
 - 2: gb_ba2:*
 - 3: gb_ba3:*
 - 4: gb_in1:*
 - 5: gb_in2:*
 - 6: gb_in3:*
 - 7: gb_om:*
 - 8: gb_ov:*
 - 9: gb_pat1:*
 - 10: gb_pat2:*
 - 11: gb_ph:*
 - 12: gb_pl1:*
 - 13: gb_pl2:*
 - 14: gb_pl3:*
 - 15: gb_pl4:*
 - 16: em_ba1:*
 - 17: em_ba2:*
 - 18: em_fun:*
 - 19: em_htgo_hum:*
 - 20: em_htgo_inv:*
 - 21: em_htgo_rod:*
 - 22: em_htg_hum1:*
 - 23: em_htg_hum2:*
 - 24: em_htg_hum3:*
 - 25: em_htg_hum4:*
 - 26: em_htg_hum5:*
 - 27: em_htg_hum6:*
 - 28: em_htg_hum7:*
 - 29: em_htg_hum8:*
 - 30: em_htg_inv1:*
 - 31: em_htg_inv2:*
 - 32: em_htg_other:*
 - 33: em_htg_rod:*
 - 34: em_humi:*
 - 35: em_hum2:*
 - 36: em_hum3:*
 - 37: em_hum4:*
 - 38: em_hum5:*
 - 39: em_hum6:*
 - 40: em_hum7:*
 - 41: em_in:*
 - 42: em_om:*
 - 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vl:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vil:*
- 59: gb_vil2:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rol:*
- 95: gb_rod:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.2	66.0	20	9	AR100356
C 2	13.2	66.0	47	9	AR032663 Sequence
C 3	13.2	66.0	47	10	I29403
C 4	13.2	66.0	47	10	I91077
C 5	12.8	64.0	30	10	I19158
6	12.8	64.0	47	9	AR032673
7	12.8	64.0	47	10	I29413
8	12.8	64.0	47	10	I91087

c 9 12.6 63.0 28 9 AR066591 Sequence
 c 10 12.6 63.0 28 10 E37410 Recombinant
 c 11 12.6 63.0 28 10 I06858 Sequence 2
 c 12 12.6 63.0 32 9 AR080756 Sequence 2
 c 13 12.6 63.0 32 10 I28948 Sequence 2
 c 14 12.6 63.0 33 10 I43985 Sequence 76
 c 15 12.6 63.0 42 97 S80775 Sequence 76
 c 16 12.4 62.0 21 10 AX096203 Sequence
 c 17 12.4 62.0 30 9 AR100568 Sequence 22
 c 18 12.4 62.0 30 10 I19159 Sequence 22
 c 19 12.2 61.0 30 10 E15705 Sequence
 c 20 12.2 61.0 34 9 AX003737 Sequence
 c 21 12.2 61.0 34 9 AX023341 Sequence
 c 22 12.2 61.0 16 9 AX061521 Sequence
 c 23 12.2 60.0 29 9 A23243 Sequence
 c 24 12.2 60.0 29 9 A23244 Sequence
 c 25 12.2 60.0 29 9 AR040875 Sequence
 c 26 12.2 60.0 29 9 AR040876 Sequence
 c 27 12.2 60.0 29 9 AR068407 Sequence
 c 28 12.2 60.0 29 9 AR068408 Sequence
 c 29 12.2 60.0 38 9 AR075847 Sequence
 c 30 12.2 60.0 39 9 A51622 Sequence 32
 c 31 12.2 60.0 40 10 I62865 Sequence 11
 c 32 12.2 60.0 49 10 E08692 Sequence
 c 33 11.8 59.0 19 9 AR021123 Sequence
 c 34 11.8 59.0 19 9 AR036267 Sequence
 c 35 11.8 59.0 21 9 AR037078 Sequence
 c 36 11.8 59.0 22 9 AR101978 Sequence
 c 37 11.8 59.0 23 10 I67849 Sequence 5
 c 38 11.8 59.0 26 10 I43986 Sequence 77
 c 39 11.8 59.0 31 9 AX001457 Sequence
 c 40 11.8 59.0 33 9 AR063468 Sequence
 c 41 11.8 59.0 33 10 I28148 Sequence 5
 c 42 11.8 59.0 40 9 AX027275 Sequence
 c 43 11.8 59.0 46 94 MMV56RNA
 c 44 11.8 59.0 48 9 AR050776 Sequence
 c 45 11.8 59.0 48 9 AR050777 Sequence

ALIGNMENTS

RESULT 1
 AR100356
 LOCUS AR100356 20 bp DNA PAT 14-FEB-2001
 DEFINITION Sequence 87 from patent US 6080580.
 ACCESSION AR100356
 VERSION AR100356.1 GI:12810804
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 20)
 AUTHORS Baker B.F., Bennett, C. Frank, Butler, M.M. and Shanahan, W.R. Jr.
 TITLE Antisense oligonucleotide modulation of tumor necrosis factor- α . (TNF- α .) expression
 JOURNAL Patent: US 6080580-A 87 27-JUN-2000;
 FEATURES Location/Qualifiers
 source 1..20
 BASE COUNT 5 a 3 c 9 g 3 t
 ORIGIN
 Query Match 66.0%; Score 13.2; DB 9; Length 20;
 Best Local Similarity 83.3%; Pred. No. 5.8e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 3 aggggaggtcagtgtg 20
 ||||| |||||
 Db 1 AGAGGAGAGTCAGTGTGG 18

RESULT 2
 AR032663/c
 LOCUS AR032663 47 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 275 from patent US 5869241.
 ACCESSION AR032663
 VERSION AR032663.1 GI:5948268
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 47)
 AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.
 TITLE Method of determining DNA sequence preference of a DNA-binding molecule
 JOURNAL Patent: US 5869241-A 275 09-FEB-1999;
 FEATURES Location/Qualifiers
 source 1..47
 BASE COUNT 12 a 15 c 13 g 7 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 47;
 Best Local Similarity 83.3%; Pred. No. 4.8e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 taggggaggtcagtgtg 19
 ||||| ||||| |||||
 Db 22 TAGGGCAGGCTCTGTTG 5

RESULT 3
 I29403/c
 LOCUS I29403 47 bp DNA PAT 06-FEB-1997
 DEFINITION Sequence 275 from patent US 5578444.
 ACCESSION I29403
 VERSION I29403.1 GI:1820194
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 47)
 AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M., Turin, L.M. and Fry, K.E.
 TITLE Sequence-directed DNA-binding molecules compositions and methods
 JOURNAL Patent: US 5578444-A 275 26-NOV-1996;
 FEATURES Location/Qualifiers
 source 1..47
 BASE COUNT 12 a 15 c 13 g 7 t
 ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 47;
 Best Local Similarity 83.3%; Pred. No. 4.8e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 taggggaggtcagtgtg 19
 ||||| ||||| |||||
 Db 22 TAGGGCAGGCTCTGTTG 5

RESULT 4
 I91077/c
 LOCUS I91077 47 bp DNA PAT 01-DEC-1998
 DEFINITION Sequence 275 from patent US 5726014.
 ACCESSION I91077
 VERSION I91077.1 GI:3935547
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 47)
 AUTHORS Edwards, C.A., Cantor, C.R., Andrews, B.M. and Turin, L.M.

TITLE Screening assay for the detection of DNA-binding molecules
JOURNAL Patent: US 5726014-A 275 10-MAR-1998;
FEATURES Location/Qualifiers

source
1. .47
/organism="unknown"

BASE COUNT 12 a 15 c 13 g 7 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 47;
Best Local Similarity 83.3%; Pred. No. 4.8e+04;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 tagggaggctcagtggtg 19

||||| ||||| |||||

Db 22 TAGGCCAGGCTCTGTTG 5

RESULT 5

I19158/c

LOCUS

119158 30 bp DNA PAT 07-OCT-1996

DEFINITION Sequence 21 from patent US 5502176.

ACCESSION I19158

VERSION I19158.1

KEYWORDS GI:1599513

SOURCE

Unknow.

REFERENCE 1 (bases 1 to 30)

AUTHORS Tenen,D.G.; Pahl,H.L. and Burn,T.C.

TITLE Myeloid cell specific promoter

JOURNAL Patent: US 5502176-A 21 26-MAR-1996;

FEATURES Location/Qualifiers

source 1. .30 /organism="unknown"

BASE COUNT 2 a 13 c 5 g 10 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 10; Length 30;

Best Local Similarity 87.5%; Pred. No. 8.4e+04;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 ggggaggctcagtggtg 19

||||| ||||| |||||

Db 18 GCGAGGCTCAGTGAG 3

RESULT 6

AR032673

LOCUS

AR032673 47 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 285 from patent US 5869241.

ACCESSION AR032673

VERSION AR032673.1

KEYWORDS GI:5948278

SOURCE

Unknow.

REFERENCE 1 (bases 1 to 47)

AUTHORS Edwards,C.A.; Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.

TITLE Method of determining DNA sequence preference of a DNA-binding

molecule

JOURNAL Patent: US 5869241-A 285 09-FEB-1999;

FEATURES Location/Qualifiers

source 1. .47 /organism="unknown"

BASE COUNT 14 a 15 c 12 g 6 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 9; Length 47;

Best Local Similarity 87.5%; Pred. No. 7.6e+04;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agggaggctcagtggt 18

||||| ||||| |||||

Db 8 AGGAAGGCTCAGTAT 23

RESULT 7

I29413

LOCUS

I29413 47 bp DNA PAT 06-FEB-1997

DEFINITION Sequence 285 from patent US 5578444.

ACCESSION I29413

VERSION I29413.1

KEYWORDS GI:1820204

SOURCE

Unknow.

REFERENCE 1 (bases 1 to 47)

AUTHORS Edwards,C.A.; Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.

TITLE Sequence-directed DNA-binding molecules compositions and methods

JOURNAL Patent: US 5578444-A 285 26-NOV-1996;

FEATURES Location/Qualifiers

source 1. .47 /organism="unknown"

BASE COUNT 14 a 15 c 12 g 6 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 10; Length 47;

Best Local Similarity 87.5%; Pred. No. 7.6e+04;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agggaggctcagtggt 18

||||| ||||| |||||

Db 8 AGGAAGGCTCAGTAT 23

RESULT 8

I91087

LOCUS

I91087 47 bp DNA PAT 01-DEC-1998

DEFINITION Sequence 285 from patent US 5726014.

ACCESSION I91087

VERSION I91087.1

KEYWORDS GI:3935557

SOURCE

Unknow.

REFERENCE 1 (bases 1 to 47)

AUTHORS Edwards,C.A.; Cantor,C.R., Andrews,B.M. and Turin,L.M.

TITLE Screening assay for the detection of DNA-binding molecules

JOURNAL Patent: US 5726014-A 285 10-MAR-1998;

FEATURES Location/Qualifiers

source 1. .47 /organism="unknown"

BASE COUNT 14 a 15 c 12 g 6 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 10; Length 47;

Best Local Similarity 87.5%; Pred. No. 7.6e+04;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agggaggctcagtggt 18

||||| ||||| |||||

Db 8 AGGAAGGCTCAGTAT 23

RESULT 9

AR066591/c

LOCUS

AR066591 28 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 2 from patent US 5851529.

ACCESSION AR066591

VERSION AR066591.1

KEYWORDS GI:5997813

JOURNAL Patent: US 5576202-A 2 19-NOV-1996;

FEATURES Location/Qualifiers
source 1. .32

BASE COUNT 8 a 8 c 9 g 7 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 10; Length 32;

Best Local Similarity 78.9%; Pred. No. 1e+05;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctaggggaggctcagtggtg 19

||||| ||| ||||| ||

Db 29 CTAGGTGAGACTCATTTGAG 11

RESULT 14

I43985/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

Query Match 63.0%; Score 12.6; DB 10; Length 33;

Best Local Similarity 78.9%; Pred. No. 1e+05;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tagggggaggctcagtggtg 20

||||| ||| ||||| ||

Db 21 TATGGGAGCTCAGTTAG 3

RESULT 15

S80775

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REMARK

FEATURES

source

BASE COUNT

ORIGIN

/db_xref="taxon:9606"

1. .42

/partial

/gene="gamma delta T cell antigen receptor delta-chain"

10 a 9 c 17 g 6 t

BASE COUNT

ORIGIN

Query Match 63.0%; Score 12.6; DB 97; Length 42;

Best Local Similarity 78.9%; Pred. No. 9.9e+04;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctaggggaggctcagtggtg 19

||||| ||| ||||| ||

Db 14 CTGGGGGAGCCCCAGTGAG 32

Search completed: October 2, 2001, 15:56:49

Job time: 14173 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:04 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-16

Perfect score: 20

Sequence: 1 ctaggggaggtccagtgtgg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: em_esthum29:*
169: em_esthum30:*
170: em_esthum31:*
171: em_esthum32:*
172: em_esthum33:*
173: em_esthum34:*
174: em_esthum35:*
175: em_esthum36:*
176: em_esthum37:*
177: em_esthum38:*
178: em_esthum39:*
179: em_esthum40:*
180: em_esthum41:*
181: em_esthum42:*
182: em_esthum43:*
183: em_esthum44:*
184: em_esthum45:*
185: em_esthum46:*
186: em_esthum47:*
187: em_esthum48:*
188: em_esthum49:*
189: em_esthum50:*

```

```

190: gb_estl10:*
191: gb_estl11:*
192: gb_estl12:*
193: gb_estl13:*
194: gb_estl14:*
195: gb_estl15:*
196: gb_estl16:*
197: gb_estl17:*
198: gb_estl18:*
199: gb_estl19:*
200: gb_estl20:*
201: gb_estl21:*
202: gb_estl22:*
203: gb_estl23:*
204: gb_estl24:*
205: gb_estl25:*
206: gb_estl26:*
207: gb_estl27:*
208: gb_estl28:*
209: gb_estl29:*
210: gb_estl30:*
211: gb_estl31:*
212: gb_estl32:*
213: gb_estl33:*
214: gb_estl34:*
215: gb_estl35:*
216: gb_estl36:*
217: gb_estl37:*
218: gb_estl38:*
219: gb_estl39:*
220: gb_estl40:*
221: gb_estl41:*
222: gb_estl42:*
223: gb_estl43:*
224: gb_estl44:*
225: gb_estl45:*
226: gb_estl46:*
227: gb_estl47:*
228: gb_estl48:*
229: gb_estl49:*
230: gb_estl50:*
231: gb_estl51:*
232: gb_estl52:*
233: gb_estl53:*
234: gb_estl54:*
235: gb_estl55:*
236: gb_estl56:*
237: gb_estl57:*
238: gb_estl58:*
239: gb_estl59:*
240: gb_estl60:*
241: gb_estl61:*
242: gb_estl62:*
243: gb_estl63:*
244: gb_estl64:*
245: gb_estl65:*
246: gb_estl66:*
247: gb_estl67:*
248: gb_estl68:*
249: gb_estl69:*
250: gb_estl70:*
251: gb_estl71:*
252: gb_estl72:*
253: gb_estl73:*
254: gb_estl74:*
255: gb_estl75:*
256: gb_estl76:*
257: gb_estl77:*
258: gb_estl78:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

JOURNAL
COMMENT
mRNAs using full-length enriched and 5'-end enriched cDNA libraries
Unpublished (2001)
Contact: Yutaka Suzuki

SUMMARIES

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	13.2	66.0	50	107	AU107335	AU107335 AU107335
2	13.2	66.0	50	107	AU107336	AU107336 AU107336
3	12.6	63.0	33	250	AZ826713	AZ826713 2M0102H12
4	12.6	63.0	46	102	AI809473	AI809473 Wn76C10.x
5	12	60.0	42	245	AZ498822	AZ498822 LM0336117
6	12	60.0	48	250	AZ833039	AZ833039 2M0113P15
7	11.8	59.0	21	247	AZ660221	AZ660221 LM0538M09
8	11.8	59.0	50	107	AU104075	AU104075 AU104075
9	11.8	59.0	50	107	AU105603	AU105603 AU105603
10	11.8	59.0	50	107	AU105604	AU105604 AU105604
11	11.8	59.0	50	107	AU105605	AU105605 AU105605
12	11.8	59.0	50	107	AU105606	AU105606 AU105606
13	11.6	58.0	20	243	AZ417155	AZ417155 LM192P07
14	11.6	58.0	26	258	TA91803P	TA91803P T. brucei
15	11.6	58.0	40	189	T95881	T95881 y647b02.sl
16	11.6	58.0	42	167	BE386700	BE386700 601274107
17	11.6	58.0	43	13	AA923260	AA923260 cm48H1.s
18	11.6	58.0	46	19	AI360975	AI360975 qy02f12.x
19	11.6	58.0	47	139	BE738401	BE738401 601572411
20	11.6	58.0	48	139	BE738353	BE738353 601572748
21	11.6	58.0	48	139	BE738724	BE738724 601572972
22	11.6	58.0	50	107	AU106034	AU106034 AU106034
23	11.4	57.0	31	14	AI000235	AI000235 ov10802.s
24	11.4	57.0	36	244	AZ484337	AZ484337 LM0310C14
25	11.4	57.0	40	246	AZ597058	AZ597058 LM0410M08
26	11.4	57.0	44	250	AZ818897	AZ818897 2M0089N10
27	11.2	56.0	34	250	AZ844396	AZ844396 2M0143B22
28	11.2	56.0	30	250	AZ815328	AZ815328 2M0083B03
29	11.2	56.0	37	242	AZ387829	AZ387829 LM0147F19
30	11.2	56.0	41	258	TA116D11P	TA116D11P T. brucei
31	11.2	56.0	42	246	AZ593675	AZ593675 LM0403G15
32	11.2	56.0	43	4	AA238745	AA238745 mx82f05.r
33	11.2	56.0	46	159	AZ24850	AZ24850 yx70e02.sl
34	11	55.0	22	246	AZ623308	AZ623308 LM0460R21
35	11	55.0	35	258	TA177A10Q	TA177A10Q T. brucei
36	11	55.0	37	13	AA933928	AA933928 on94f04.s
37	11	55.0	39	247	AZ659784	AZ659784 LM0537M24
38	11	55.0	41	244	AZ470576	AZ470576 LM0284B14
39	11	55.0	42	243	AZ440678	AZ440678 LM0231C19
40	11	55.0	43	247	AZ660216	AZ660216 LM0538L07
41	11	55.0	44	158	H92874	H92874 yf91g11.sl
42	11	55.0	49	241	AZ313961	AZ313961 LM0030I02
43	11	55.0	49	258	TA307F09Q	TA488871 T. brucei
44	11	55.0	50	107	AU106627	AU106627 AU106627
45	10.8	54.0	19	249	AZ775541	AZ775541 2M0008H17

ALIGNMENTS

```

FEATURES
source
    1. 50
    Location/Qualifiers
        organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="LNG11142"
        /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT
    6 a 11 c 25 g 8 t

```

Query Match	66.0%	Score 13.2;	DB 107;	Length 50;
Best Local Similarity	83.3%	Pred. No. 3.7e+04;		
Matches 15; Conservative		0; Mismatches 3;	Indels 0;	Gaps 0;
QY	1	ctaggggaggctcagtg	18	
Db	32	CTAGGGGGCGGCTCGAGT	49	

ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 50)
AUTHORS	Suzuki, X., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo, K., Suyama, A. and Sugano, S.
TITLE	Fine structural analysis of transcription start sites of human

```

Db 11 CTAGGTGGCGCAGTG 29

RESULT 4
LOCUS AI809473 46 bp mRNA EST 20-DEC-1999
DEFINITION wh76c10.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2386674 3'
similar to TR:Q40402 Q40402 EXTENSIN ;, mRNA sequence.
ACCESSION AI809473
VERSION AI809473.1 GI:53966039
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 46)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 782 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES
source
1..46
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2386674"
/clone_lib="NCI_CGAP_CLL1"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/notes="vector: p7T3b-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGCGCGCATGTGCTTTTGTGTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 14 a 25 c 1 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 102; Length 46;
Best Local Similarity 78.9%; Pred. No. 7.1e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 taggggaggtcagtgtg 20
||||| | |||| |
Db 46 TAGGGGTTGTTTCAGGGTGG 28

RESULT 5
LOCUS AZ498822 42 bp DNA GSS 05-OCT-2000
DEFINITION lM0336117F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0336117 F, DNA sequence.
ACCESSION AZ498822
VERSION AZ498822.1 GI:10677034
KEYWORDS GSS.

Mus musculus
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 33)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0102 row: H column: 12
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 33.

FEATURES
source
1..33
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0102H12"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g114732114|g1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 6 a 8 c 13 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 250; Length 33;
Best Local Similarity 78.9%; Pred. No. 7e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctaggggaggtccagtgtg 19
||||| | || |||||

```

```

Db 11 CTAGGTGGCGCAGTG 29

RESULT 4
LOCUS AI809473 46 bp mRNA EST 20-DEC-1999
DEFINITION wh76c10.x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2386674 3'
similar to TR:Q40402 Q40402 EXTENSIN ;, mRNA sequence.
ACCESSION AI809473
VERSION AI809473.1 GI:53966039
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 46)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 782 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES
source
1..46
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2386674"
/clone_lib="NCI_CGAP_CLL1"
/tissue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/notes="vector: p7T3b-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGGCGCGCATGTGCTTTTGTGTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 14 a 25 c 1 g 6 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 102; Length 46;
Best Local Similarity 78.9%; Pred. No. 7.1e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 taggggaggtcagtgtg 20
||||| | |||| |
Db 46 TAGGGGTTGTTTCAGGGTGG 28

RESULT 5
LOCUS AZ498822 42 bp DNA GSS 05-OCT-2000
DEFINITION lM0336117F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0336117 F, DNA sequence.
ACCESSION AZ498822
VERSION AZ498822.1 GI:10677034
KEYWORDS GSS.

```

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 42)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0336 row: 1 column: 17
 Seq primer: CGTGTAAACGAGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 42.
FEATURES
 Location/Qualifiers
 1..42
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWB42 (gil14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 8 a 2 c 25 g 7 t
ORIGIN
 Query Match 60.0%; Score 12; DB 245; Length 42;
 Best Local Similarity 75.0%; Pred. No. 1.3e+05;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 ctaggggaggctcagtgtg 20
 |||||
 Db 20 CTAGGGGAGCTGGGGGGG 39
RESULT 6
LOCUS AZ833039 48 bp DNA GSS 20-FEB-2001
DEFINITION 2M0113P15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0113P15 R, DNA sequence.
ACCESSION AZ833039
VERSION AZ833039.1 GI:13002947

KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 48)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0113 row: P column: 15
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 48.
FEATURES
 Location/Qualifiers
 1..48
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWB42 (gil14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 16 a 8 c 17 g 7 t
ORIGIN
 Query Match 50.0%; Score 12; DB 250; Length 48;
 Best Local Similarity 75.0%; Pred. No. 1.4e+05;
 Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 ctaggggaggctcagtgtg 20
 |||||
 Db 7 CTGGGGGAGGGGCGTGGGG 26
RESULT 7
LOCUS AZ660221/c 21 bp DNA GSS 14-DEC-2000
DEFINITION IM0538M09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0538M09 F, DNA sequence.
ACCESSION AZ660221

```

VERSION      AZ660221.1  GI:11797367
KEYWORDS     GSS.
SOURCE       Mus musculus
ORGANISM     Mus musculus

REFERENCE    1 (bases 1 to 21)
AUTHORS      Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
             Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
             ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
             and Wright,D., Weiss,R.
TITLE        Mouse whole genome scaffolding with paired end reads from 10kb
             plasmid inserts
JOURNAL      Unpublished (2000)
COMMENT      Contact: Robert B. Weiss
             University of Utah Genome Center
             Rm. 309, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
             84112, USA
             Tel: 801 585 5606
             Fax: 801 585 7177
             Email: dunn@genetics.utah.edu
             Insert Length: 10000 Std Error: 0.00
             Plate: 0538 row: M column: 09
             Seq primer: CATTGTAACGACGCCAGT
             Class: plasmid inserts
             High quality sequence stop: 21.
             Location/Qualifiers
                 1..21
                 /organism="Mus musculus"
                 /strain="C57BL/6J"
                 /db_xref="taxon:10090"
                 /clone="UUGC1M0538M09"
                 /clone.lib="Mouse 10kb plasmid UUGC1M library"
                 /sex="Male"
                 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
                 /note="Vector: PWD42hv; Purified genomic DNA from M.
                 musculus C57BL/6J (male) was obtained from the Jackson
                 Laboratory Mouse DNA Resource
                 (http://www.jax.org/resources/documents/dnares/). The DNA
                 was hydrodynamically sheared by repeated passage through a
                 0.005 inch orifice at constant velocity. The sheared DNA
                 was blunt end-repaired with T4 DNA polymerase and T4
                 polynucleotide kinase. Adaptor oligonucleotides were
                 ligated to the blunt ends in high molar excess. The
                 adaptor DNA was purified and size-selected for a 9.5 to
                 10.5 kb range using preparative agarose gel
                 electrophoresis. Vector DNA was prepared from a derivative
                 of PWD42 (g114732114[gblAF129072.1]), a copy-number
                 inducible derivative of plasmid R1. The vector was ligated
                 with adaptors complementary to the insert adaptors and
                 purified. The sheared, adaptor mouse DNA was annealed to
                 adaptor vector DNA, and transformed into
                 chemically-competent E. coli XL10-Gold (Stratagene) cells
                 and selected for ampicillin resistance."
BASE COUNT   7 a 11 c 1 g 2 t
ORIGIN

Query Match      59.0%; Score 11.8; DB 247; Length 21;
Best Local Similarity 86.7%; Pred. No. 1.6e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 gggaggctcagtg 19
   ||| ||| ||| |||
Db 16 GGGTGGTCAGTGTG 2

RESULT 8
AUI04075      50 bp mRNA EST 05-APR-2001
LOCUS         AUI04075 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION    HEP14513, mRNA sequence.
ACCESSION     AUI04075
VERSION       AUI04075.1 GI:13553596
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 50)
AUTHORS       H., Ota,T., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
             ,K., Suyama,A. and Sugano,S.
TITLE         Fine Structural analysis of transcription start sites of human
             mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL       Unpublished (2001)
COMMENT       Contact: Yutaka Suzuki
             Department of Virology
             Institute of Medical Science, University of Tokyo
             4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
             Email: ysuzuki@ims.u-tokyo.ac.jp
             Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
             ,S. Construction and characterization of a full length-enriched and
             a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
             Location/Qualifiers
                 1..50
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="HEP14513"
                 /clone.lib="Sugano Homo sapiens cDNA library"
BASE COUNT     8 a 11 c 20 g 11 t
ORIGIN

Query Match      59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 1.7e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agggaggctcagtg 17
   ||| ||| ||| |||
Db 17 AGGGAGGCGCGGTG 31

RESULT 9
AUI05603/c    50 bp mRNA EST 05-APR-2001
LOCUS         AUI05603 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION    HSI00510, mRNA sequence.
ACCESSION     AUI05603
VERSION       AUI05603.1 GI:13555124
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 50)
AUTHORS       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE         Fine Structural analysis of transcription start sites of human
             mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL       Unpublished (2001)
COMMENT       Contact: Yutaka Suzuki
             Department of Virology
             Institute of Medical Science, University of Tokyo
             4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
             Email: ysuzuki@ims.u-tokyo.ac.jp
             Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
             ,S. Construction and characterization of a full length-enriched and
             a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
             Location/Qualifiers
                 1..50
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="HSI00510"
BASE COUNT     7 a 11 c 1 g 2 t
ORIGIN

Query Match      59.0%; Score 11.8; DB 247; Length 21;
Best Local Similarity 86.7%; Pred. No. 1.6e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 gggaggctcagtg 19
   ||| ||| ||| |||
Db 16 GGGTGGTCAGTGTG 2

RESULT 8
AUI04075      50 bp mRNA EST 05-APR-2001
LOCUS         AUI04075 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
DEFINITION    HEP14513, mRNA sequence.
ACCESSION     AUI04075
VERSION       AUI04075.1 GI:13553596
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 50)
AUTHORS       H., Ota,T., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
             ,K., Suyama,A. and Sugano,S.
TITLE         Fine Structural analysis of transcription start sites of human
             mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL       Unpublished (2001)
COMMENT       Contact: Yutaka Suzuki
             Department of Virology
             Institute of Medical Science, University of Tokyo
             4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
             Email: ysuzuki@ims.u-tokyo.ac.jp
             Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
             ,S. Construction and characterization of a full length-enriched and
             a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
             Location/Qualifiers
                 1..50
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="HEP14513"
BASE COUNT     8 a 11 c 20 g 11 t
ORIGIN

```


BASE COUNT 9 a 18 c 11 g 12 t
ORIGIN /clone_lib="Sugano Homo sapiens cDNA library"

Query Match 59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 1.7e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 gggaggctcagtgtg 19
||||| ||||| |||||
Db 41 GGGAGCCTCATTGTG 27

RESULT 10
AU105604/c

LOCUS AU105604 50 bp mRNA EST 05-APR-2001
DEFINITION AU105604 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HS101710, mRNA sequence.

ACCESSION AU105604
VERSION AU105604.1 GI:13555125
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.

TITLE Fine structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL Unpublished (2001)

COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yszukie@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source

1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HS101710"

BASE COUNT 10 a 18 c 8 g 14 t
ORIGIN /clone_lib="Sugano Homo sapiens cDNA library"

Query Match 59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 1.7e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 gggaggctcagtgtg 19
||||| ||||| |||||
Db 44 GGGAGCCTCATTGTG 30

RESULT 11
AU105605/c

LOCUS AU105605 50 bp mRNA EST 05-APR-2001
DEFINITION AU105605 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HS101558, mRNA sequence.

ACCESSION AU105605
VERSION AU105605.1 GI:13555126
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS

Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.

TITLE

Fine structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL

Unpublished (2001)

COMMENT

Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yszukie@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HS101558"
/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 10 a 18 c 9 g 13 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 1.7e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 gggaggctcagtgtg 19
||||| ||||| |||||

Db 42 GGGAGCCTCATTGTG 28

RESULT 12
AU105606/c

LOCUS AU105606 50 bp mRNA EST 05-APR-2001
DEFINITION AU105606 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HS1016872, mRNA sequence.

ACCESSION AU105606

VERSION AU105606.1 GI:13555127

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.

TITLE Fine structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL

Unpublished (2001)

COMMENT

Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yszukie@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source

1. .50
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HS1016872"
/clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 10 a 18 c 10 g 12 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 107; Length 50;
Best Local Similarity 86.7%; Pred. No. 1.7e+05;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 5 gggaggctcagtgtg 19
    ||||| ||||| |||||
Db 40 GGGAGCTCATTTGTG 26

RESULT 13
LOCUS      A2417155/c      20 bp      DNA      GSS      03-OCT-2000
DEFINITION LM0192P07R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
            clone UUGCLM0192P07 R, DNA sequence.
ACCESSION  A2417155
VERSION     A2417155.1  GI:10541168
KEYWORDS    GSS.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 20)
AUTHORS     Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D., Weiss,R.
TITLE       Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
JOURNAL     Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0192 row: P column: 07
            Seq primer: CACACAGGAACACGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 20.

FEATURES             source
1..20
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0192P07"
/clone.lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42hiv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT      8 a      7 c      2 g      3 t
ORIGIN

Query Match      58.0%; Score 11.6; DB 243; Length 20;
Best Local Similarity 77.8%; Pred. No. 2e+05;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tagggaggctcagtgtg 19
    ||||| ||||| |||||
Db 20 TAGGGGTAGCTCTTTGTG 3

RESULT 14
LOCUS      TA91B03P      26 bp      DNA      GSS      13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 91b03, forward sequence,
            genomic survey sequence.
ACCESSION  AL459576
VERSION     AL459576.1  GI:11861682
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
ORGANISM    Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1 (bases 1 to 26)
AUTHORS     Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE       Direct Submission
JOURNAL     Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb). The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects. In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES             source
1..26
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="91b03"
BASE COUNT      6 a      8 c      8 g      4 t
ORIGIN

Query Match      58.0%; Score 11.6; DB 258; Length 26;
Best Local Similarity 77.8%; Pred. No. 2e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctaggggaggctcagtgt 18
    | |||| |||| ||||
Db 2 CAAGGGCAGGATCGGTGT 19

RESULT 15
LOCUS      T95881/c      40 bp      mRNA      EST      27-MAR-1995
DEFINITION ye47b02.si Soares fetal liver spleen INFLS Homo sapiens cDNA clone
            IMAGE:120843 3' similar to gb|M87923|HUMALCE12 Human carcinoma
            cell-derived Alu RNA transcript, (rRNA); gb:M57627 INTERLEUKIN-10
            PRECURSOR (HUMAN); mRNA sequence.
ACCESSION  T95881
VERSION     T95881.1  GI:734505
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Query Match      58.0%; Score 11.6; DB 243; Length 20;
Best Local Similarity 77.8%; Pred. No. 2e+05;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:49 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-17

Perfect score: 20

Sequence: 1 attctcgtctgtgtggaa 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba1.*

2: gb_ba2.*

3: gb_ba3.*

4: gb_in1.*

5: gb_in2.*

6: gb_in3.*

7: gb_om.*

8: gb_ov.*

9: gb_pat1.*

10: gb_pat2.*

11: gb_ph.*

12: gb_pl1.*

13: gb_pl2.*

14: gb_pl3.*

15: gb_pl4.*

16: em_ba1.*

17: em_ba2.*

18: em_fun.*

19: em_htgo_hum.*

20: em_htgo_inv.*

21: em_htgo_rod.*

22: em_htg_hum1.*

23: em_htg_hum2.*

24: em_htg_hum3.*

25: em_htg_hum4.*

26: em_htg_hum5.*

27: em_htg_hum6.*

28: em_htg_hum7.*

29: em_htg_hum8.*

30: em_htg_inv1.*

31: em_htg_inv2.*

32: em_htg_other.*

33: em_htg_rod.*

34: em_hum1.*

35: em_hum2.*

36: em_hum3.*

37: em_hum4.*

38: em_hum5.*

39: em_hum6.*

40: em_hum7.*

41: em_in.*

42: em_om.*

43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vil2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_rol.*
95: gb_rol2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
c 1	15.2	76.0	40	9 AR012105	AR012105 Sequence
c 2	15.2	76.0	40	9 AR014544	AR014544 Sequence
c 3	14.2	71.0	20	9 AR059098	AR059098 Sequence
4	13.8	69.0	46	97 HUMTCIGC1A	M96133 Human hybri
5	13.6	68.0	32	9 A73580	A73580 Sequence 4
c 6	13.4	67.0	18	9 AR067419	AR067419 Sequence
c 7	13.4	67.0	26	9 AR070795	AR070795 Sequence
c 8	13.4	67.0	39	93 HSCD85706	Z50819 H.sapiens m

c 9 13 65.0 23 9 AX044085 Sequence
 c 10 13 65.0 23 9 AX044138 Sequence
 c 11 13 65.0 23 9 AX044178 Sequence
 c 12 13 65.0 48 92 HSA224214
 c 13 12.6 63.0 20 9 AR103323
 c 14 12.6 63.0 30 10 I25076
 c 15 12.6 63.0 30 10 I71129
 c 16 12.6 63.0 37 9 A82991
 c 17 12.6 63.0 37 9 A82993
 c 18 12.6 63.0 21 9 AX027268
 c 19 12.2 61.0 20 10 I73108
 c 20 12.2 61.0 24 9 AR018080
 c 21 12.2 61.0 24 9 AR051920
 c 22 12.2 61.0 29 9 AR049382
 c 23 12.2 61.0 29 9 AR095543
 c 24 12.2 61.0 30 9 A83263
 c 25 12.2 61.0 36 93 HSAJ5143
 c 26 12.2 61.0 38 9 AR108637
 c 27 12.2 61.0 42 9 AR058955
 c 28 12.2 61.0 42 9 AR105229
 c 29 12.2 61.0 49 10 AX098901
 c 30 12.2 61.0 49 10 AX099321
 c 31 12 60.0 20 10 E26600
 c 32 12 60.0 26 10 E04774
 c 33 12 60.0 27 9 AR090975
 c 34 12 60.0 28 9 AR050042
 c 35 12 60.0 28 9 AR064551
 c 36 12 60.0 31 9 AR013956
 c 37 12 60.0 31 10 I23438
 c 38 12 60.0 31 10 I25985
 c 39 12 60.0 33 9 A68960
 c 40 12 60.0 45 9 AR065636
 c 41 12 60.0 45 9 AR068600
 c 42 12 60.0 45 9 AR095952
 c 43 12 60.0 45 10 I23529
 c 44 11.8 59.0 18 9 AR078900
 c 45 11.8 59.0 23 9 AR106304

ALIGNMENTS

RESULT 1
 AR012105/c
 LOCUS AR012105 40 bp DNA PAT 04-DEC-1998
 DEFINITION Sequence 52 from patent US 5763209.
 ACCESSION AR012105
 VERSION AR012105.1 GI:3970095
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 40)
 AUTHORS Sukhtame,V.P.
 TITLE Methods and materials relating to the functional domains of DNA binding proteins
 JOURNAL Patent: US 5763209-A 52 09-JUN-1998;
 FEATURES Location/Qualifiers
 source 1..40
 BASE COUNT 9 a 14 c 13 g 4 t
 ORIGIN
 Query Match 76.0%; Score 15.2; DB 9; Length 40;
 Best Local Similarity 85.0%; Pred. No. 4.6e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 attctcgtcgtcgtggtgaa 20
 ||||||| ||||| |||||
 Db 21 ATTCTCGGTCTGCGCGAA 2
 ||||||| ||||| |||||
 RESULT 2
 AR014544/c
 LOCUS AR014544 40 bp DNA PAT 05-DEC-1998
 DEFINITION Sequence 52 from patent US 5773583.
 ACCESSION AR014544
 VERSION AR014544.1 GI:3971998
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 40)
 AUTHORS Sukhtame,V.P.
 TITLE Methods and materials relating to the functional domains of DNA binding proteins
 JOURNAL Patent: US 5773583-A 52 30-JUN-1998;
 FEATURES Location/Qualifiers
 source 1..40
 BASE COUNT 9 a 14 c 13 g 4 t
 ORIGIN
 Query Match 76.0%; Score 15.2; DB 9; Length 40;
 Best Local Similarity 85.0%; Pred. No. 4.6e+04;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 attctcgtcgtcgtggtgaa 20
 ||||||| ||||| |||||
 Db 21 ATTCTCGGTCTGCGCGAA 2
 ||||||| ||||| |||||
 RESULT 3
 AR059098/c
 LOCUS AR059098 20 bp DNA PAT 29-SEP-1999
 DEFINITION Sequence 16 from patent US 5837854.
 ACCESSION AR059098
 VERSION AR059098.1 GI:5984675
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 20)
 AUTHORS Mulder,C.
 TITLE Oligonucleotides with anti-Epstein-Barr virus activity
 JOURNAL Patent: US 5837854-A 16 17-NOV-1998;
 FEATURES Location/Qualifiers
 source 1..20
 BASE COUNT 5 a 7 c 4 g 4 t
 ORIGIN
 Query Match 71.0%; Score 14.2; DB 9; Length 20;
 Best Local Similarity 84.2%; Pred. No. 1.5e+05;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 2 ttctcgtcgtcgtggtgaa 20
 ||||||| ||||| |||||
 Db 20 TTGCTCGCAGCTGGAGAA 2
 ||||||| ||||| |||||
 RESULT 4
 HUMTCIGC1A
 LOCUS HUMTCIGC1A 46 bp DNA PRI 13-JAN-1995
 DEFINITION Human hybrid T cell receptor/immunoglobulin protein, clone 1 V-D-J segment.
 ACCESSION M96133
 VERSION M96133.1 GI:339188
 KEYWORDS T-cell receptor/immunoglobulin protein.
 SOURCE Homo sapiens DNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 46)
 AUTHORS Kobayashi,Y., Tycko,B., Soreng,A.L. and Sklar,J.
 TITLE Transrearrangements between antigen receptor genes in normal human lymphoid tissues and in ataxia telangiectasia
 JOURNAL J. Immunol. 147 (9), 3201-3209 (1991)
 MEDLINE 92013154
 FEATURES Location/Qualifiers

source
 1..46
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene
 1..3
 /gene="IGHV"
 /map="14q32.33"
 V_region
 1..3
 /gene="IGHV"
 /note="G00-128-528"
 gene
 4..13
 /gene="IGH"
 /map="14q32.33"
 N_region
 4..13
 /gene="IGH"
 /note="G00-118-731"
 D_segment
 14..18
 /gene="TCRD"
 /note="G00-120-406"
 gene
 14..46
 /gene="TCRD"
 /map="14q11.2"
 N_region
 19..23
 /gene="TCRD"
 /note="G00-120-406"
 D_segment
 24..30
 /gene="TCRD"
 /note="G00-120-406"
 N_region
 31..35
 /gene="TCRD"
 /note="G00-120-406"
 J_segment
 36..46
 /gene="TCRD"
 /note="G00-120-406"
 BASE COUNT 9 a 16 c 11 g 10 t
 ORIGIN

Query Match 69.0%; Score 13.8; DB 97; Length 46;
 Best Local Similarity 88.2%; Pred. No. 1.8e+05;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ttctcgtcgtcgtggtg 18
 |||||
 Db 14 TTCCTCGCTCCTGGGG 30

RESULT 5
 A73580
 LOCUS A73580 32 bp DNA
 DEFINITION Sequence 4 from Patent WO9418959.
 ACCESSION A73580
 VERSION A73580.1 GI:6064168
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unidentified
 unclassified.

REFERENCE 1 (bases 1 to 32)
 AUTHORS Nemeth,E.F. and Brown,E.M.
 TITLE CALCIUM RECEPTOR-ACTIVE MOLECULES
 JOURNAL Patent: WO 9418959-A 4 01-SEP-1994;
 BRIGHAM & WOMENS HOSPITAL (US); NPS PHARMA INC (US)
 FEATURES Location/Qualifiers

source
 1..32
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 5 a 4 c 8 g 8 t 7 others

ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 32;
 Best Local Similarity 70.0%; Pred. No. 2.4e+05;
 Matches 14; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 attctcgtcgtcgtgtaa 20
 |||||
 Db 5 ACTNCTGGTCTGGTGGAR 24

RESULT 6
 AR067419/c
 LOCUS AR067419 18 bp DNA
 DEFINITION Sequence 16 from patent US 5851763.
 ACCESSION AR067419
 VERSION AR067419.1 GI:5998641
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 18)
 AUTHORS Heym,B., Cole,S., Young,D., Zhang,Y., Honore,N., Telenti,A. and Bodmer,T.
 TITLE Rapid detection of antibiotic resistance in mycobacterium tuberculosis
 JOURNAL Patent: US 5851763-A 16 22-DEC-1998;
 FEATURES Location/Qualifiers

source
 1..18
 /organism="unknown"
 BASE COUNT 5 a 7 c 5 g 1 t
 ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 18;
 Best Local Similarity 93.3%; Pred. No. 3.5e+05;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ctgcgtcgtcgtgga 19
 |||||
 Db 16 CTCGCTGGGTGGA 2

RESULT 7
 AR070795/c
 LOCUS AR070795 26 bp DNA
 DEFINITION Sequence 2 from patent US 5908764.
 ACCESSION AR070795
 VERSION AR070795.1 GI:7221683
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.

REFERENCE 1 (bases 1 to 26)
 AUTHORS Brunker,P., Minas,W., Kallio,P. and Bailey,J.E.
 TITLE Methods and compositions for increasing production of erythromycin
 JOURNAL Patent: US 5908764-A 2 01-JUN-1999;
 FEATURES Location/Qualifiers

source
 1..26
 /organism="unknown"
 BASE COUNT 6 a 9 c 3 g 8 t
 ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 26;
 Best Local Similarity 93.3%; Pred. No. 3.1e+05;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 tcgctcgtcgtgtaa 20
 |||||
 Db 26 TAGCTGCTGGTGAA 12

```

RESULT 8
HSCD85706/c 39 bp mRNA PRI 15-MAR-1996
LOCUS H.sapiens mRNA for T cell receptor (ID C85706).
DEFINITION Z50819
ACCESSION Z50819.1 GI:1004241
VERSION constant region; joining region; T cell receptor; variable region.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 39)
AUTHORS Wang,E.C., Moss,P.A., Frodsham,P., Lehner,P.J., Bell,J.I. and
Borysiewicz,L.K.
TITLE Oligoclonal CD8+, CD57+ T cells in normal, healthy individuals and
their responses to HCMV
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 39)
AUTHORS Wang,E.C.
TITLE Direct Submission
JOURNAL Submitted (17-AUG-1995) Eddie CY Wang, Medicine, University of
Wales of College of, Medicine, Heath Park, Cardiff, South
Glamorgan, CF4 4XN, Wales UK
REFERENCE 3 (bases 1 to 39)
AUTHORS Wang,E.C., Moss,P.A., Frodsham,P., Lehner,P.J., Bell,J.I. and
Borysiewicz,L.K.
TITLE CD8highCD57+ T lymphocytes in normal, healthy individuals are
oligoclonal and respond to human cytomegalovirus
JOURNAL J. Immunol. 155 (10), 5046-5056 (1995)
MEDLINE 96062329
FEATURES
source Location/Qualifiers
1..39
/organism="Homo sapiens"
/isolate="EW"
/db_xref="taxon:9606"
1..12
/note="end T cell receptor V beta 24"
13..24
/note="CDR3"
25..39
/note="beginning J beta 2.5"
BASE COUNT 9 a 11 c 15 g 4 t
ORIGIN
Query Match 67.0%; Score 13.4; DB 93; Length 39;
Best Local Similarity 93.3%; Pred. No. 2.7e+05;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 cctcgctgctggtg 18
|||||
Db 19 CCTCTCTGCTGCTGG 5

RESULT 9
AX044085/c 23 bp DNA PAT 24-NOV-2000
LOCUS AX044085
DEFINITION Sequence 40 from Patent WO0066748.
ACCESSION AX044085
VERSION AX044085.1 GI:11342963
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 23)
AUTHORS Hawkes,T.R., Warner,S.A., Andrews,C.J., Bachoo,S. and
Pickerill,A.P.
TITLE Herbicide resistant plants
JOURNAL Patent: WO 0066748-A 40 09-NOV-2000;
ZENECA LIMITED (GB)
FEATURES
source Location/Qualifiers
1..23
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer"

```

```

/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer"
5 a 7 c 5 g 6 t

BASE COUNT 5 a 7 c 5 g 6 t
ORIGIN
Query Match 65.0%; Score 13; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.8e+05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 gctgctggtggaa 20
|||||
Db 20 GCTGCTGCTGGAA 8

RESULT 10
AX044138/c 23 bp DNA PAT 24-NOV-2000
LOCUS AX044138
DEFINITION Sequence 38 from Patent WO0066747.
ACCESSION AX044138
VERSION AX044138.1 GI:11343016
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 23)
AUTHORS Hawkes,T.R., Warner,S.A., Andrews,C.J., Bachoo,S. and
Pickerill,A.P.
TITLE Herbicide resistant plants
JOURNAL Patent: WO 0066747-A 38 09-NOV-2000;
ZENECA LIMITED (GB)
FEATURES
source Location/Qualifiers
1..23
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer"
5 a 7 c 5 g 6 t

BASE COUNT 5 a 7 c 5 g 6 t
ORIGIN
Query Match 65.0%; Score 13; DB 9; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.8e+05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 gctgctggtggaa 20
|||||
Db 20 GCTGCTGCTGGAA 8

RESULT 11
AX044178/c 23 bp DNA PAT 24-NOV-2000
LOCUS AX044178
DEFINITION Sequence 30 from Patent WO0066746.
ACCESSION AX044178
VERSION AX044178.1 GI:11343056
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 23)
AUTHORS Hawkes,T.R., Warner,S.A., Andrews,C.J., Bachoo,S. and
Pickerill,A.P.
TITLE Herbicide resistant plants
JOURNAL Patent: WO 0066746-A 30 09-NOV-2000;
ZENECA LIMITED (GB)
FEATURES
source Location/Qualifiers
1..23
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Primer"
5 a 7 c 5 g 6 t

BASE COUNT 5 a 7 c 5 g 6 t
ORIGIN

```


Db 1 TTCATCGCTGCTGCGCA 19

Search completed: October 2, 2001, 15:56:50
Job time: 14174 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:06 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-17

Perfect score: 20

Sequence: 1 attctcgtcgtctggtggaa 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_est4.*
- 5: gb_est5.*
- 6: gb_est6.*
- 7: gb_est7.*
- 8: gb_est8.*
- 9: gb_est9.*
- 10: gb_est10.*
- 11: gb_est11.*
- 12: gb_est12.*
- 13: gb_est13.*
- 14: gb_est14.*
- 15: gb_est15.*
- 16: gb_est16.*
- 17: gb_est17.*
- 18: gb_est18.*
- 19: gb_est19.*
- 20: gb_est20.*
- 21: gb_est21.*
- 22: gb_est22.*
- 23: gb_est23.*
- 24: gb_est24.*
- 25: gb_est25.*
- 26: gb_est26.*
- 27: gb_est27.*
- 28: gb_est28.*
- 29: gb_est29.*
- 30: gb_est30.*
- 31: gb_est31.*
- 32: gb_est32.*
- 33: gb_est33.*
- 34: gb_est34.*
- 35: gb_est35.*
- 36: gb_est36.*
- 37: gb_est37.*
- 38: gb_est38.*
- 39: gb_est39.*
- 40: gb_est40.*
- 41: gb_est41.*
- 42: gb_est42.*
- 43: gb_est43.*
- 44: gb_est44.*
- 45: gb_est45.*
- 46: gb_est46.*
- 47: gb_est47.*

- 44: em_esthum10.*
- 45: em_esthum11.*
- 46: em_esthum12.*
- 47: em_esthum13.*
- 48: em_esthum14.*
- 49: em_esthum15.*
- 50: em_esthum16.*
- 51: em_esthum17.*
- 52: em_esthum18.*
- 53: em_esthum19.*
- 54: em_esthum20.*
- 55: em_esthum21.*
- 56: em_esthum22.*
- 57: em_esthum23.*
- 58: em_esthum24.*
- 59: em_esthum25.*
- 60: em_esthum26.*
- 61: em_esthum27.*
- 62: em_esthum28.*
- 63: em_estin1.*
- 64: em_estin2.*
- 65: em_estin3.*
- 66: em_estin4.*
- 67: em_estin5.*
- 68: em_estom1.*
- 69: em_estom2.*
- 70: em_estov1.*
- 71: em_estov2.*
- 72: em_estpl1.*
- 73: em_estpl2.*
- 74: em_estpl3.*
- 75: em_estpl4.*
- 76: em_estpl5.*
- 77: em_estpl6.*
- 78: em_estpl7.*
- 79: em_estpl8.*
- 80: em_estpl9.*
- 81: em_estpl10.*
- 82: em_estro1.*
- 83: em_estro2.*
- 84: em_estro3.*
- 85: em_estro4.*
- 86: em_estro5.*
- 87: em_estro6.*
- 88: em_estro7.*
- 89: em_estro8.*
- 90: em_estro9.*
- 91: em_estro10.*
- 92: em_estro11.*
- 93: em_estro12.*
- 94: em_estro13.*
- 95: em_estro14.*
- 96: em_estro15.*
- 97: em_estro16.*
- 98: em_estro17.*
- 99: em_estro18.*
- 100: em_estro19.*
- 101: em_estro20.*
- 102: gb_est25.*
- 103: gb_est26.*
- 104: gb_est27.*
- 105: gb_est28.*
- 106: gb_est29.*
- 107: gb_est30.*
- 108: gb_est31.*
- 109: gb_est32.*
- 110: gb_est41.*
- 111: gb_est42.*
- 112: gb_est43.*
- 113: gb_est44.*
- 114: gb_est45.*
- 115: gb_est46.*
- 116: gb_est47.*

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*
 183: gb_est114:*
 184: gb_est115:*
 185: gb_est116:*
 186: gb_est117:*
 187: gb_est118:*
 188: gb_est119:*
 189: gb_est120:*

190: gb_est121:*
 191: gb_est122:*
 192: gb_est123:*
 193: gb_est124:*
 194: gb_est125:*
 195: gb_est126:*
 196: gb_est127:*
 197: gb_est128:*
 198: gb_est129:*
 199: gb_est130:*
 200: gb_est131:*
 201: gb_est132:*
 202: gb_est133:*
 203: gb_est134:*
 204: gb_est135:*
 205: gb_est136:*
 206: gb_est137:*
 207: gb_est138:*
 208: gb_est139:*
 209: gb_est140:*
 210: gb_est141:*
 211: gb_est142:*
 212: gb_est143:*
 213: gb_est144:*
 214: gb_est145:*
 215: gb_est146:*
 216: gb_est147:*
 217: gb_est148:*
 218: gb_est149:*
 219: gb_est150:*
 220: gb_est151:*
 221: gb_est152:*
 222: gb_est153:*
 223: gb_est154:*
 224: gb_est155:*
 225: gb_est156:*
 226: gb_est157:*
 227: gb_est158:*
 228: gb_est159:*
 229: gb_est160:*
 230: gb_est161:*
 231: gb_est162:*
 232: gb_est163:*
 233: gb_est164:*
 234: gb_est165:*
 235: gb_est166:*
 236: gb_est167:*
 237: gb_est168:*
 238: gb_est169:*
 239: gb_est170:*
 240: gb_est171:*
 241: gb_est172:*
 242: gb_est173:*
 243: gb_est174:*
 244: gb_est175:*
 245: gb_est176:*
 246: gb_est177:*
 247: gb_est178:*
 248: gb_est179:*
 249: gb_est180:*
 250: gb_est181:*
 251: gb_est182:*
 252: gb_est183:*
 253: gb_est184:*
 254: gb_est185:*
 255: gb_est186:*
 256: gb_est187:*
 257: gb_est188:*
 258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.


```

COMMENT      Contact: Robert B. Weiss
              University of Utah
              Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
              84112, USA
              Tel: 801 585 5606
              Fax: 801 585 7177
              Email: ddunn@genetics.utah.edu
              Insert Length: 10000 Std Error: 0.00
              Plate: 0288 row: 1 column: 14
              Seq primer: CACACAGGAACAGCTATGACC
              Class: plasmid ends
              High quality sequence stop: 32.
              Location/Qualifiers
                1..32
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUCG1M0288I14"
                /clone_lib="Mouse 10kb plasmid UUCG1M library"
                /sex="Male"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                /note="vector: PWD42nv; Purified genomic DNA from M.
                musculus C57BL/6J (male) was obtained from the Jackson
                Laboratory Mouse DNA Resource
                (http://www.jax.org/resources/documents/dnares/). The DNA
                was hydrodynamically sheared by repeated passage through a
                0.005 inch orifice at constant velocity. The sheared DNA
                was blunt end-repaired with T4 DNA polymerase and T4
                polynucleotide kinase. Adaptor oligonucleotides were
                ligated to the blunt ends in high molar excess. The
                adapted DNA was purified and size-selected for a 9.5 to
                10.5 kb range using preparative agarose gel
                electrophoresis. Vector DNA was prepared from a derivative
                of PWD42 (gi14732114[gb]AF129072.1), a copy-number
                inducible derivative of plasmid R1. The vector was ligated
                with adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
                7 a      8 c      10 g      7 t

BASE COUNT
ORIGIN

Query Match      66.0%; Score 13.2; DB 244; Length 32;
Best Local Similarity 83.3%; Pred. No. 3.9e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 attctcgtcgtggtgg 18
   | | | | | | | | | |
Db 15 ACTCCTCTGCTGCTGG 32

RESULT 3
AUI07725/c
LOCUS      AUI07725      50 bp      mRNA      EST      05-APR-2001
DEFINITION AUI07725 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
BX3A0030, mRNA sequence.
ACCESSION  AUI07725
VERSION     AUI07725.1 GI:13557246
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 50)
AUTHORS    Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
            Fine structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
            Unpublished (2001)
JOURNAL
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
            1..50
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="BX3A0030"
            /clone_lib="Sugano Homo sapiens cDNA library"
            8 a      19 c      15 g      7 t      1 others

BASE COUNT
ORIGIN

Query Match      65.0%; Score 13; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.9e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 cgctgctggtgga 19
   | | | | | | | | | |
Db 21 CGCTGCTGCTGGA 9

RESULT 4
AUI07726/c
LOCUS      AUI07726      50 bp      mRNA      EST      05-APR-2001
DEFINITION AUI07726 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
COL02114, mRNA sequence.
ACCESSION  AUI07726
VERSION     AUI07726.1 GI:13557247
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 50)
AUTHORS    Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
            ,K., Suyama,A. and Sugano,S.
            Fine structural analysis of transcription start sites of human
            mRNAs using full-length enriched and 5'-end enriched cDNA libraries
            Unpublished (2001)
JOURNAL
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
            1..50
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="COL02114"
            /clone_lib="Sugano Homo sapiens cDNA library"
            7 a      21 c      14 g      8 t

BASE COUNT
ORIGIN

Query Match      65.0%; Score 13; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.9e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 cgctgctggtgga 19
   | | | | | | | | | |
Db 43 CGCTGCTGCTGGA 31

JOURNAL
COMMENT     Unpublished (2001)

```

```

RESULT 5
AUI07729/c
LOCUS      50 bp      mRNA      EST      05-APR-2001
DEFINITION Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP01448, mRNA sequence.
ACCESSION AUI07729
VERSION    AUI07729.1 GI:13557250
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
          ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
          ,K., Suyama,A. and Sugano,S.
TITLE     Fine Structural analysis of transcription start sites of human
          mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL   Unpublished (2001)
COMMENT   Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          ,S. Construction and characterization of a full length-enriched and
          a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             Location/Qualifiers
     source            1..50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP01448"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      8 a      21 c      13 g      8 t
ORIGIN
Query Match      65.0%; Score 13; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.9e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cgctgctggtgga 19
|||||
DB 49 CGCTGCTGGTGGG 37

RESULT 6
AUI07730/c
LOCUS      50 bp      mRNA      EST      05-APR-2001
DEFINITION Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP01701, mRNA sequence.
ACCESSION AUI07730
VERSION    AUI07730.1 GI:13557251
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
          ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
          ,K., Suyama,A. and Sugano,S.
TITLE     Fine Structural analysis of transcription start sites of human
          mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL   Unpublished (2001)
COMMENT   Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          ,S. Construction and characterization of a full length-enriched and
          a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

```

```

FEATURES             Location/Qualifiers
     source            1..50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP01701"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      8 a      23 c      13 g      6 t
ORIGIN
Query Match      65.0%; Score 13; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.9e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cgctgctggtgga 19
|||||
DB 44 CGCTGCTGGTGGG 32

RESULT 7
AUI07732/c
LOCUS      50 bp      mRNA      EST      05-APR-2001
DEFINITION Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP13686, mRNA sequence.
ACCESSION AUI07732
VERSION    AUI07732.1 GI:13557253
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS   Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
          ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
          ,K., Suyama,A. and Sugano,S.
TITLE     Fine Structural analysis of transcription start sites of human
          mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL   Unpublished (2001)
COMMENT   Contact: Yutaka Suzuki
          Department of Virology
          Institute of Medical Science, University of Tokyo
          4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
          Email: ysuzuki@ims.u-tokyo.ac.jp
          ,S. Construction and characterization of a full length-enriched and
          a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES             Location/Qualifiers
     source            1..50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP13686"
                     /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT      8 a      22 c      14 g      6 t
ORIGIN
Query Match      65.0%; Score 13; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.9e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cgctgctggtgga 19
|||||
DB 43 CGCTGCTGGTGGG 31

RESULT 8
AUI07733/c
LOCUS      50 bp      mRNA      EST      05-APR-2001
DEFINITION Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP13813, mRNA sequence.
ACCESSION AUI07733
VERSION    AUI07733.1 GI:13557254
KEYWORDS   EST.

```

```

SOURCE
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
               ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
               ,K., Suyama,A. and Sugano,S.
TITLE        Fine Structural analysis of transcription start sites of human
               mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL
COMMENT      Unpublished (2001)
               Contact: Yutaka Suzuki
               Department of Virology
               Institute of Medical Science, University of Tokyo
               4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
               Email: yszuki@iems.u-tokyo.ac.jp
               Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
               ,S. Construction and characterization of a full length-enriched and
               a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
source       1..50
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="HEP13813"
               /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT   8 a 23 c 13 g 6 t
ORIGIN
Query Match 65.0%; Score 13; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.9e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 cgcctcgtgtgga 19
   |||||||
Db 45 CCCTCGTGTGGA 33

RESULT 9
AUI07734/c
LOCUS      AUI07734      50 bp      mRNA      EST      05-APR-2001
DEFINITION AUI07734 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
HEP14963, mRNA sequence.
ACCESSION  AUI07734
VERSION    AUI07734.1 GI:13557255
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 50)
AUTHORS    Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
               ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
               ,K., Suyama,A. and Sugano,S.
TITLE      Fine Structural analysis of transcription start sites of human
               mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL
COMMENT    Unpublished (2001)
               Contact: Yutaka Suzuki
               Department of Virology
               Institute of Medical Science, University of Tokyo
               4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
               Email: yszuki@iems.u-tokyo.ac.jp
               Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
               ,S. Construction and characterization of a full length-enriched and
               a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
source       1..50
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="HEP14963"
               /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT   8 a 20 c 13 g 9 t
ORIGIN

```

```

Query Match 65.0%; Score 13; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.9e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 cgcctcgtgtgga 19
   |||||||
Db 43 CGCTGCTGTGGA 31

RESULT 10
AUI07736/c
LOCUS      AUI07736      50 bp      mRNA      EST      05-APR-2001
DEFINITION AUI07736 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
KAT09254, mRNA sequence.
ACCESSION  AUI07736
VERSION    AUI07736.1 GI:13557257
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 50)
AUTHORS    Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
               ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
               ,K., Suyama,A. and Sugano,S.
TITLE      Fine Structural analysis of transcription start sites of human
               mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL
COMMENT    Unpublished (2001)
               Contact: Yutaka Suzuki
               Department of Virology
               Institute of Medical Science, University of Tokyo
               4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
               Email: yszuki@iems.u-tokyo.ac.jp
               Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
               ,S. Construction and characterization of a full length-enriched and
               a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
source       1..50
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="KAT09254"
               /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT   7 a 20 c 14 g 9 t
ORIGIN
Query Match 65.0%; Score 13; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.9e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 cgcctcgtgtgga 19
   |||||||
Db 39 CGCTGCTGTGGA 27

RESULT 11
AUI07737/c
LOCUS      AUI07737      50 bp      mRNA      EST      05-APR-2001
DEFINITION AUI07737 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
REC00765, mRNA sequence.
ACCESSION  AUI07737
VERSION    AUI07737.1 GI:13557258
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 50)
AUTHORS    Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
               ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
               ,K., Suyama,A. and Sugano,S.

```


TITLE Fine Structural analysis of transcription start sites of human
JOURNAL mRNAs using full-length enriched and 5'-end enriched cDNA libraries
COMMENT Unpublished (2001)
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, A. and Sugano
 S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source
 1. .50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="REC00765"
 /clone_lib="Sugano Homo sapiens cDNA library"
 8 a 22 c 14 g 6 t

BASE COUNT
ORIGIN

Query Match 65.0%; Score 13; DB 107; Length 50;
 Best Local Similarity 100.0%; Pred. No. 4.9e+04;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cgctgctggtgga 19
 |||||
 Db 43 CGCTGCTGGTGA 31

RESULT 12
AU107738/c

LOCUS AU107738 50 bp mRNA EST 05-APR-2001
DEFINITION AU107738 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 2RV61429, mRNA sequence.

ACCESSION AU107738
VERSION AU107738.1 GI:13557259
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata
 H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo
 K., Suyama, A. and Sugano, S.

TITLE Fine Structural analysis of transcription start sites of human
JOURNAL mRNAs using full-length enriched and 5'-end enriched cDNA libraries
COMMENT Unpublished (2001)
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, A. and Sugano
 S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

source
 1. .50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2RV61429"
 /clone_lib="Sugano Homo sapiens cDNA library"
 7 a 22 c 15 g 6 t

BASE COUNT
ORIGIN

Query Match 65.0%; Score 13; DB 107; Length 50;
 Best Local Similarity 100.0%; Pred. No. 4.9e+04;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cgctgctggtgga 19
 |||||

Db 39 CGCTGCTGGTGA 27

RESULT 13
AU107739/c

LOCUS AU107739 50 bp mRNA EST 05-APR-2001
DEFINITION AU107739 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 2RV61978, mRNA sequence.

ACCESSION AU107739
VERSION AU107739.1 GI:13557260
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki, Y., Tsunoda, T., Taira, H., Mizushima-Sugano, J., Sese, J., Hata
 H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y., Morishita, S., Okubo
 K., Suyama, A. and Sugano, S.

TITLE Fine Structural analysis of transcription start sites of human
JOURNAL mRNAs using full-length enriched and 5'-end enriched cDNA libraries
COMMENT Unpublished (2001)
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp

Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
 S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source
 1. .50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="2RV61978"
 /clone_lib="Sugano Homo sapiens cDNA library"
 7 a 22 c 15 g 6 t

BASE COUNT
ORIGIN

Query Match 65.0%; Score 13; DB 107; Length 50;
 Best Local Similarity 100.0%; Pred. No. 4.9e+04;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 cgctgctggtgga 19
 |||||
 Db 39 CGCTGCTGGTGA 27

RESULT 14
AZ599531

LOCUS AZ599531 45 bp DNA GSS 13-DEC-2000
DEFINITION IM0414M23R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0414M23 R, DNA sequence.

ACCESSION AZ599531
VERSION AZ599531.1 GI:11721721
KEYWORDS GSS.
SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 45)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL plasmid inserts
COMMENT Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0414 row: M column: 23

Seq primer: CACACAGGAACACGCTATGACC

Class: plasmid ends

High quality sequence stop: 45.

FEATURES

Location/Qualifiers

source

1..45

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGCIM0414M23"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PWD42 (gil4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT

6 a 10 c 8 g 21 t

ORIGIN

Query Match 64.0%; Score 12.8; DB 246; Length 45;
 Best Local Similarity 87.5%; Pred. No. 6.1e+04;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ttctctcgctgctggtg 17

||||| ||||| |||||

Db 20 TTCCTGCTGCTGCTG 35

RESULT 15

AU105903

LOCUS

AU105903 50 bp mRNA EST 05-APR-2001

DEFINITION AU105903 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone

KAT08022, mRNA sequence.

ACCESSION AU105903

VERSION AU105903.1 GI:13555424

KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 50)

Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata

,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo

,K., Suyama,A. and Sugano,S.

Fine Structural analysis of transcription start sites of human

mRNAs using full-length enriched and 5'-end enriched cDNA libraries

Unpublished (2001)

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: ysuzuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 ,S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

Location/Qualifiers

source

1..50

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="KAT08022"

/clone_lib="Sugano Homo sapiens cDNA library"

3 a 21 c 9 g 17 t

Query Match 64.0%;

Best Local Similarity 87.5%; Score 12.8; DB 107; Length 50;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ttctctcgctgctggtg 17

||||| ||||| |||||

Db 13 TTCGTCGCTGCTGCTG 28

Search completed: October 2, 2001, 15:01:07

Job time: 10836 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:51 ; Search time 3339.34 Seconds
(without alignments)
92.640 Million cell updates/sec

Title: US-09-757-100B-20
Perfect score: 20
Sequence: 1 catgatgcttaaaagcttac 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_hum1_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vil:*
59: gb_vil2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_rol2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
C 1	13.8	69.0	29	9 AR038139	AR038139 Sequence
C 2	13.8	69.0	29	10 I79867	I79867 Sequence 7
C 3	13.6	68.0	48	9 A87694	A87694 Sequence 2
C 4	13.4	67.0	44	9 AR075944	AR075944 Sequence
C 5	13.2	66.0	20	45 E10769	E10769 PCR primer
C 6	13.2	66.0	20	45 E10798	E10798 PCR primer
C 7	13.2	66.0	33	9 A97414	A97414 Sequence 4
C 8	13.2	66.0	33	9 AR022050	AR022050 Sequence

```
c 9 13.2 66.0 33 10 I92840 Sequence 14
c 10 12.6 63.0 35 9 A61915 Sequence 3
c 11 12.2 61.0 25 9 AR049871 Sequence 3
c 12 12.2 61.0 25 10 I59435 Sequence 10
c 13 12.2 61.0 27 9 AR110319 Sequence 10
c 14 12.2 61.0 29 9 A65301 Sequence 24
c 15 12.2 61.0 29 9 AR036134 Sequence
c 16 12.2 61.0 29 9 AR069509 Sequence
c 17 12.2 61.0 29 9 AR076177 Sequence
c 18 12.2 61.0 30 10 I77055 Sequence 2
c 19 12.2 61.0 40 9 A65366 Sequence 2
c 20 12 60.0 30 10 AX100542 Sequence
c 21 12 60.0 44 9 AR071481 Sequence
c 22 12 60.0 44 9 AR080501 Sequence
c 23 11.8 59.0 21 9 AR071089 Sequence
c 24 11.8 59.0 30 9 AR063799 Sequence
c 25 11.8 59.0 30 10 I30162 Sequence 84
c 26 11.8 59.0 34 9 AR079838 Sequence
c 27 11.8 59.0 34 53 C75779 Homo sapien
c 28 11.8 59.0 37 9 AR028960 Sequence
c 29 11.8 59.0 45 9 A23344 Artificial
c 30 11.8 59.0 45 9 AR068522 Sequence
c 31 11.8 59.0 47 54 G54572 AffySTS0000
c 32 11.6 58.0 21 10 E27054 Novel recep
c 33 11.6 58.0 23 9 AR079557 Sequence
c 34 11.6 58.0 25 9 AR075352 Sequence
c 35 11.6 58.0 26 9 A48469 Sequence 8
c 36 11.6 58.0 26 9 AR050991 Sequence
c 37 11.6 58.0 26 9 AR088908 Sequence
c 38 11.6 58.0 26 9 AX000672 Sequence
c 39 11.6 58.0 30 10 I23896 Sequence 13
c 40 11.6 58.0 32 10 I17561 Sequence 24
c 41 11.6 58.0 33 10 AX084116 Sequence
c 42 11.6 58.0 33 10 E15443 PCR primer
c 43 11.6 58.0 35 9 A21585 XDH3 oligon
c 44 11.6 58.0 38 9 AR084169 Sequence
c 45 11.6 58.0 40 9 A36486 Sequence 27
```

ALIGNMENTS

```
RESULT 1
AR038139/c
LOCUS AR038139 29 bp DNA
DEFINITION Sequence 9 from patent US 5804417.
ACCESSION AR038139
VERSION AR038139.1 GI:5956856
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 29)
Martens,G.JulianusMaria, Chaudhuri,B. and Stephan,C.
TITLE Recombinant production of proteins using 7B2 protein
JOURNAL Patent: US 5804417-A 9 08-SEP-1998;
METHOD OF PREPARING AND APPLICATIONS
LOCATION/Qualifiers
source 1. .29
BASE COUNT 11 a 6 c 3 g 9 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 9; Length 29;
Best Local Similarity 88.2%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 tgatgcttaaaagctta 19
||| ||| ||| ||| ||| |||
Db 18 TGTTCATATAAAGCTTA 2

RESULT 2
AR075944
LOCUS AR075944 44 bp DNA
DEFINITION Sequence 5 from patent US 5958713.
ACCESSION AR075944
VERSION AR075944.1 GI:10002690
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 48)
Drevet,P. and Menez,A.
TITLE NON-COVALENT COMPLEX COMPRISING AT LEAST AN ANTIBODY AND ELEMENT
BINDING WITH IMMUNOGLOBULINS ASSOCIATED WITH AN ACTIVE SUBSTANCE,
METHOD OF PREPARING AND APPLICATIONS
PATENT: WO 9834956-A 2 13-AUG-1998;
COMMISSARIAT ENERGIE ATOMIQUE (FR); DREVET PASCAL (FR)
LOCATION/Qualifiers
source 1. .48
BASE COUNT 14 a 9 c 10 g 15 t
ORIGIN
Query Match 68.0%; Score 13.6; DB 9; Length 48;
Best Local Similarity 80.0%; Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 catgatgcttaaaagcttac 20
||| ||| ||| ||| ||| |||
Db 8 CAGGATCCTTATAAGCTTCC 27

RESULT 4
AR075944
LOCUS AR075944 44 bp DNA
DEFINITION Sequence 5 from patent US 5958713.
ACCESSION AR075944
VERSION AR075944.1 GI:10002690
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 29)
Martens,G.JulianusMaria, Chaudhuri,B. and Stephan,C.
TITLE Recombinant production of proteins using 7B2 protein
JOURNAL Patent: US 5804417-A 7 13-JAN-1998;
METHOD OF PREPARING AND APPLICATIONS
LOCATION/Qualifiers
source 1. .29
BASE COUNT 11 a 6 c 3 g 9 t
ORIGIN
Query Match 69.0%; Score 13.8; DB 9; Length 29;
Best Local Similarity 88.2%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 tgatgcttaaaagctta 19
||| ||| ||| ||| ||| |||
Db 18 TGTTCATATAAAGCTTA 2
```

REFERENCE 1 (bases 1 to 44)
 AUTHORS Thastrup,O., Tullin,S. slashed.ren, Poulsen,L.Kongsbak and Bj.o slashed.rm,S.petersen.
 TITLE Method of detecting biologically active substances by using green fluorescent protein
 JOURNAL Patent: US 5958713-A 5 28-SEP-1999;
 FEATURES Location/Qualifiers
 source 1..44
 BASE COUNT 12 a 8 c 11 g 13 t
 ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 44;
 Best Local Similarity 93.3%; Pred. No. 2.6e+04;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 atgcttaaaagctta 19
 || |||||
 Db 5 ATTCTTAAAGCTTA 19

RESULT 5
 E10769/c standard; DNA; UNC; 20 BP.
 ID E10769
 XX
 AC E10769;
 XX
 SV E10769.1

08-OCT-1997 (Rel. 52, Created)
 02-SEP-2000 (Rel. 65, Last updated, Version 2)
 PCR primer for gaining human serum albumin.
 XX JP 1996051982-A/11.

OS None
 OC Artificial sequences.
 ON JP 1996051982-A/11
 XX unclassified.
 XX

[1]
 1-20
 RA Higashida H., Murakami K., Hama Y., Tsukamoto Y., Isoai A., Kumagai H.;
 RT "MODIFIED GENE CODING HUMAN SERUM ALBUMEN";
 RL Patent number JP1996051982-A/11, 27-FEB-1996.
 RL ASAHI GLASS CO LTD.
 XX

OS None
 CC Artificial sequences.
 CC PN JP 1996051982-A/11
 CC PD 27-FEB-1996
 CC PF 11-AUG-1994 JP 1994209369
 CC PI HIGASHIDA HIDEKI, MURAKAMI KIMIKO, HAMA YUKO, TSUKAMOTO YOKO,
 CC PI ISOAI ATSUSHI, KUMAGAI HIROMICHI
 CC PC C07K19/00,C12N1/19,C12P21/02,(C12N1/19,C12R1:645),
 CC PC (C12P21/02,
 CC PC C12R1:645);
 CC CC strandedness: Single;
 CC CC topology: Linear;
 CC CC hypothetical: No;
 CC CC anti-sense: Yes;
 CC FH Key Location/Qualifiers
 CC FT source 1..20 /organism="Artificial sequences"

XX
 FH Key Location/Qualifiers
 FH FT source 1..20 /db_xref="taxon:32644"
 FT /organism="unidentified"

XX
 SQ Sequence 20 BP; 6 A; 3 C; 2 G; 9 T; 0 other;
 A97414

Query Match 66.0%; Score 13.2; DB 45; Length 20;
 Best Local Similarity 83.3%; Pred. No. 3.2e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 atgatgcttaaaagctta 19
 |||||
 Db 20 ATGAAGATCAAAAGCTTA 3

RESULT 6
 E10798/c standard; DNA; UNC; 20 BP.
 ID E10798
 XX
 AC E10798;
 XX
 SV E10798.1

08-OCT-1997 (Rel. 52, Created)
 02-SEP-2000 (Rel. 65, Last updated, Version 2)
 PCR primer for gaining human serum albumin.
 XX JP 1996053500-A/8.

OS None
 OC Artificial sequences.
 ON JP 1996053500-A/8
 XX unclassified.
 XX

[1]
 1-20
 RA Higashida H., Murakami K., Hama Y., Tsukamoto Y., Isoai A., Kumagai H.;
 RT "FUSED PROTEIN AND GENE CODING SAID PROTEIN";
 RL Patent number JP1996053500-A/8, 27-FEB-1996.
 RL ASAHI GLASS CO LTD.
 XX

OS None
 CC Artificial sequences.
 CC PN JP 1996053500-A/8
 CC PD 27-FEB-1996
 CC PF 11-AUG-1994 JP 1994209368
 CC PI HIGASHIDA HIDEKI, MURAKAMI KIMIKO, HAMA YUKO, TSUKAMOTO YOKO,
 CC PI ISOAI ATSUSHI, KUMAGAI HIROMICHI
 CC PC C07K19/00,C12N1/19,C12N15/09,C12P21/02,(C12N1/19,C12R1:645),
 CC PC (C12P21/02,
 CC PC C12R1:645);
 CC CC strandedness: Single;
 CC CC topology: Linear;
 CC FH Key Location/Qualifiers
 CC FT source 1..20 /organism="Artificial sequences"

XX
 FH Key Location/Qualifiers
 FH FT source 1..20 /db_xref="taxon:32644"
 FT /organism="unidentified"

SQ Sequence 20 BP; 6 A; 3 C; 2 G; 9 T; 0 other;
 Query Match 66.0%; Score 13.2; DB 45; Length 20;
 Best Local Similarity 83.3%; Pred. No. 3.2e+04;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 atgatgcttaaaagctta 19
 |||||
 Db 20 ATGAAGATCAAAAGCTTA 3

RESULT 7
 A97414

LOCUS A97414 33 bp DNA PAT 26-JAN-2000
DEFINITION Sequence 4 from Patent WO9916867.
ACCESSION A97414
VERSION A97414.1 GI:6780760
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 33)
AUTHORS Waldmann, H. and Zelenika, D.
TITLE TRANSGENIC MODEL COMPRISING TCR ALPHA AND BETA CHAINS
JOURNAL Patent: WO 9916867-A 4 08-APR-1999;
COBOLD STEPHEN P (GB); ISIS INNOVATION (GB)
FEATURES
source
1. .33
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 10 a 5 c 9 g 9 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 33;
Best Local Similarity 83.3%; Pred. No. 3.3e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 catgatgcttaaaagctt 18
|| ||||| ||||| ||
Db 8 CAGGATGCATAAAGTTT 25

RESULT 8
AR022050/c
LOCUS AR022050 33 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 14 from patent US 5792456.
ACCESSION AR022050
VERSION AR022050.1 GI:3976112
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 33)
AUTHORS Yelton, D., Glaser, S., Huse, W. and Rosok, M. Joanne.
TITLE Mutant BR96 antibodies reactive with human carcinomas
JOURNAL Patent: US 5792456-A 14 11-AUG-1998;
FEATURES
source
1. .33
/organism="unknown"
BASE COUNT 12 a 7 c 7 g 7 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 33;
Best Local Similarity 83.3%; Pred. No. 3.3e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 tgatgcttaaaagcttac 20
| ||||| ||||| ||
Db 28 TTATGCTTAAAGCTTTC 11

RESULT 9
I92840/c
LOCUS I92840 33 bp DNA PAT 01-DEC-1998
DEFINITION Sequence 14 from patent US 5728821.
ACCESSION I92840
VERSION I92840.1 GI:3937310
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 33)
AUTHORS Yelton, D., Glaser, S., Huse, W. and Rosok, M. Joanne.
TITLE Mutant BR96 antibodies reactive with human carcinomas

JOURNAL Patent: US 5728821-A 14 17-MAR-1998;
FEATURES
source
1. .33
/organism="unknown"
BASE COUNT 12 a 7 c 7 g 7 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 10; Length 33;
Best Local Similarity 83.3%; Pred. No. 3.3e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 tgatgcttaaaagcttac 20
| ||||| ||||| ||
Db 28 TTATGCTTAAAGCTTTC 11

RESULT 10
A61915/c
LOCUS A61915 35 bp DNA PAT 09-MAR-1998
DEFINITION Sequence 3 from Patent WO9712969.
ACCESSION A61915
VERSION A61915.1 GI:3716021
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 35)
AUTHORS Poustka, A. and Coy, J.
TITLE FMRI-RELATED PROTEIN
JOURNAL Patent: WO 9712969-A 3 10-APR-1997;
DEUTSCHES KREBSFORSCH (DE)
COMMENT Other publication DE 19534763 970515.
FEATURES
source
1. .35
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 10 a 7 c 7 g 11 t
ORIGIN

Query Match 63.0%; Score 12.6; DB 9; Length 35;
Best Local Similarity 78.9%; Pred. No. 6.4e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 atgatgcttaaaagcttac 20
|| || ||||| ||||| ||
Db 20 ATAAATCTTAAAGCTTCC 2

RESULT 11
AR049871/c
LOCUS AR049871 25 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 10 from patent US 5824777.
ACCESSION AR049871
VERSION AR049871.1 GI:5971863
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 25)
AUTHORS Sasaki, K., Mori, T. and Makino, S.
TITLE Attenuated measles virus vaccine containing specific nucleotide
sequence and a method for its absolute identification
JOURNAL Patent: US 5824777-A 10 20-OCT-1998;
FEATURES
source
1. .25
/organism="unknown"
BASE COUNT 8 a 7 c 6 g 4 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 25;

Best Local Similarity 82.4%; Pred. No. 1e+05; Mismatches 0; Indels 3; Gaps 0;

QY 1 catgatgcttaaaagct 17
||||| | | | |

Db 25 CATGATGCTTGAGAGGT 9

RESULT 12
I59435/c

LOCUS I59435 25 bp DNA PAT 07-OCT-1997

DEFINITION Sequence 10 from patent US 5654136.

ACCESSION I59435

VERSION I59435.1 GI:2478067

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 25)
AUTHORS Sasaki,K., Mori,T. and Makino,S.
TITLE Attenuated measles virus vaccine, containing specific nucleotide sequence and a method for its absolute identification

JOURNAL Patent: US 5654136-A 10 05-AUG-1997;

FEATURES Location/Qualifiers
source 1..25
/organism="unknown"

BASE COUNT 8 a 7 c 6 g 4 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 10; Length 25;
Best Local Similarity 82.4%; Pred. No. 1e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 catgatgcttaaaagct 17
-||||| | | | |

Db 25 CATGATGCTTGAGAGGT 9

RESULT 13
AR110319/c

LOCUS AR110319 27 bp DNA PAT 14-FEB-2001

DEFINITION Sequence 13 from patent US 6114503.

ACCESSION AR110319

VERSION AR110319.1 GI:12826595

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 27)
AUTHORS Wei,Y., Ruben,S.M., Sancar,A., Hsu,S.D. and Kazantsev,A.G.
TITLE Human blue-light photoreceptor hCRY2

JOURNAL Patent: US 6114503-A 13 05-SEP-2000;

FEATURES Location/Qualifiers
source 1..27
/organism="unknown"

BASE COUNT 4 a 8 c 8 g 7 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 27;
Best Local Similarity 82.4%; Pred. No. 1e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 atgatgcttaaaagctt 18
| | | | | | | | | |

Db 20 AGGATGCTGAAAGCTT 4

RESULT 14
A65301

LOCUS A65301 29 bp DNA PAT 29-MAR-1999

DEFINITION Sequence 24 from patent WO9735983.

ACCESSION A65301.1 GI:4531080

VERSION A65301.1 GI:4531080

KEYWORDS

SOURCE Unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 29)
AUTHORS Jepson,I., Greenland,A.J., Thomas,D.R. and Philippe.
TITLE CYSTEINE PROTEASE PROMOTER FROM OIL SEED RAPE AND A METHOD FOR THE CONTAINMENT OF PLANT GERMPLASM

JOURNAL Patent: WO 9735983-A 24 02-OCT-1997;
ZENECA LTD (GB)

COMMENT Other publication AU 2033797 19971017.

FEATURES Location/Qualifiers
source 1..29
/organism="unidentified"
/db_xref="taxon:32644"
/clone="CYSGI2CR"

BASE COUNT 7 a 4 c 8 g 10 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 29;
Best Local Similarity 82.4%; Pred. No. 1e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 atgatgcttaaaagctt 18
| | | | | | | | | |

Db 2 ATGATGGTGACAAAGCTT 18

RESULT 15
AR036134/c

LOCUS AR036134 29 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 4 from patent US 5871991.

ACCESSION AR036134

VERSION AR036134.1 GI:5952802

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 29)
AUTHORS Elrod,S.L. and Cherry,J.R.
TITLE Aspergillus oryzae 5-aminolevulinic acid synthases and nucleic acids encoding same

JOURNAL Patent: US 5871991-A 4 16-FEB-1999;

FEATURES Location/Qualifiers
source 1..29
/organism="unknown"

BASE COUNT 7 a 8 c 8 g 6 t

ORIGIN

Query Match 61.0%; Score 12.2; DB 9; Length 29;
Best Local Similarity 82.4%; Pred. No. 1e+05;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 tgatgcttaaaagctta 19
| | | | | | | | | |

Db 18 TGCTGCTTAAATGCATA 2

Search completed: October 2, 2001, 15:56:52
Job time: 14176 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:08 ; Search time 10798.2 Seconds
(without alignments)
17.508 Million cell updates/sec

Title: US-09-757-100B-20

Perfect score: 20

Sequence: 1 catgatgcttaaaagcttac 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est25:*
- 26: gb_est26:*
- 27: gb_est27:*
- 28: gb_est28:*
- 29: gb_est29:*
- 30: gb_est30:*
- 31: gb_est31:*
- 32: gb_est32:*
- 33: gb_est33:*
- 34: gb_est34:*
- 35: gb_est35:*
- 36: gb_est36:*
- 37: gb_est37:*
- 38: gb_est38:*
- 39: gb_est39:*
- 40: gb_est40:*
- 41: gb_est41:*
- 42: gb_est42:*
- 43: gb_est43:*
- 44: gb_est44:*
- 45: gb_est45:*
- 46: gb_est46:*
- 47: gb_est47:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est33:*
- 111: gb_est34:*
- 112: gb_est35:*
- 113: gb_est36:*
- 114: gb_est37:*
- 115: gb_est38:*
- 116: gb_est39:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				TITLE	
Result No.	Score	Query Match %	Length DB ID	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts	
1	13.2	66.0	50 243 AZ403428	Unpublished (2000)	
2	13	65.0	27 249 AZ785639	Contact: Robert B. Weiss	
3	13	65.0	39 189 U74031	University of Utah	
4	12.8	64.0	37 258 TA37H11P	Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA	
5	12.2	61.0	30 250 AZ833603	Tel: 801 585 5606	
6	12.2	61.0	31 258 TA52H11Q	Fax: 801 585 7177	
7	12.2	61.0	37 15 AI024143	Email: ddunn@genetics.utah.edu	
8	12.2	61.0	48 141 BF144195	Insert Length: 10000 Std Error: 0.00	
9	12.2	61.0	48 141 BE896146	Plate: 0171 row: I column: 17	
10	11.8	59.0	37 158 HA0544	Seq primer: CGTGTAAACGACGCGCAGT	
11	11.8	59.0	48 258 TA78C03Q	Class: plasmid ends	
12	11.8	59.0	49 258 TA213B01Q	High quality sequence stop: 50.	
13	11.8	59.0	50 189 T73552	Location/Qualifiers	
14	11.6	58.0	32 258 TA318D05P	1. 50	
15	11.6	58.0	38 244 AZ487122	/organism="Mus musculus"	
16	11.6	58.0	47 242 AZ345901	/strain="C57BL/6J"	
17	11.6	58.0	50 107 AU105753	/db_xref="taxon:10090"	
18	11.4	57.0	36 249 AZ800965	/clone="UUGC1M017117"	
19	11.4	57.0	40 242 AZ378207	/clone_lib="Mouse 10kb plasmid UUGC1M library"	
20	11.4	57.0	50 258 TA154F10Q	/sex="Male"	
21	11.2	56.0	34 250 AZ819685	/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"	
22	11.2	56.0	40 8 AA508697	/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource	
23	11.2	56.0	44 247 AZ666447	(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114[gb]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor-mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."	
24	11.2	56.0	45 244 AZ451564	BASE COUNT 12 a 9 c 13 g 16 t	
25	11.2	56.0	47 152 BG324633	ORIGIN	
26	11.2	56.0	49 6 AA387830	Query Match 66.0%; Score 13.2; DB 243; Length 50;	
27	11	55.0	34 244 AZ462320	Best Local Similarity 83.3%; Pred. No. 1.7e+04;	
28	11	55.0	38 250 AZ822308	Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
29	11	55.0	41 244 AZ451720	Qy 3 tgatgcttaaaagcttac 20	
30	11	55.0	46 243 AZ437798		
31	11	55.0	46 249 AZ792431	Db 11 TTATGCTTAAATCTCAC 28	
32	11	55.0	46 258 TA97E06Q	RESULT 2	
33	11	55.0	50 107 AU103861	AZ785639 27 bp DNA GSS 16-FEB-2001	
34	11	55.0	50 156 D19972	LOCUS 2M0029012R Mouse 10kb plasmid UUGC1M library Mus musculus genomic	
35	10.8	54.0	21 246 AZ612315	DEFINITION clone UUGC1M017117 F, DNA sequence.	
36	10.8	54.0	35 249 AZ792265	ACCESSION AZ785639	
37	10.8	54.0	36 249 AZ800965	VERSION AZ785639	
38	10.8	54.0	36 250 AZ818349	KEYWORDS GSS.	
39	10.8	54.0	43 14 AI006816	SOURCE house mouse.	
40	10.8	54.0	43 241 AZ312629	ORGANISM Mus musculus	
41	10.8	54.0	43 241 AZ335682	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
42	10.8	54.0	49 244 AZ467762	1 (bases 1 to 50)	
43	10.8	54.0	50 107 AU103706	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.	
44	10.6	53.0	28 249 AZ780363	and Wright,D.,Weiss,R.	
45	10.6	53.0	32 249 AZ775467		

TITLE and Wright,D., Weiss,R.
JOURNAL Mouse whole genome scaffolding with paired end reads from 10kb
COMMENT plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: dunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Seq primer: CACACAGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 27.
FEATURES Location/Qualifiers
 1..27
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUC240029012"
 /clone_lib="Mouse 10Kb plasmid UUC1m library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g14732114|g14732114|g14732114), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 17 a 3 c 2 g 5 t
ORIGIN

Query Match 65.0%; Score 13; DB 249; Length 27;
 Best Local Similarity 100.0%; Pred. No. 2e+04;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 atgatgcttaaaa 14
 |||||
Db 5 ATGATGCTTAAAA 17

RESULT 3
LOCUS U74031/c 39 bp mRNA EST 05-SEP-1997
DEFINITION RN074031 rat lambda ZAPII library (C.P.Hamel) Rattus norvegicus
ACCESSION U74031
VERSION U74031.1 GI:1658134
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
REFERENCE 1 (bases 1 to 39)
AUTHORS Soto-Prior,A., Lavigne-Rebillard,M., Lenoir,M., Ripoll,C.,

TITLE Rebillard,G., Vago,P., Pujol,R. and Hamel,C.P.
JOURNAL Identification of preferentially expressed cochlear genes by
COMMENT systematic sequencing of a rat cochlea cDNA library
 Brain Res. Mol. Brain Res. 47 (1-2), 1-10 (1997)
 97364947
 Contact: Hamel, CP
 Laboratoire de Neurobiologie de l'Audition - INSERM U254
 Institut National de la Sante et de la Recherche Medicale
 C.H.R. St-Charles, 34295 Montpellier cedex, France
 Tel: (33) 04 67 33 69 75
 Fax: (33) 04 67 52 56 01
 Email: biomol@net.fr
 also expressed in brain, cerebellum, eye, lung, kidney and liver as
 determined by RT-PCR.

FEATURES Location/Qualifiers
 1..39
 /organism="Rattus norvegicus"
 /strain="Wistar"
 /db_xref="taxon:10116"
 /clone="pC083"
 /clone_lib="rat lambda ZAPII library (C.P.Hamel)"
 /dev_stage="postnatal day 24"
 /lab_host="Escherichia coli"
 /note="Organ: cochlea; Vector: lambda ZAPII; Site_1: EcoRI
 ; Site_2: XhoI"
BASE COUNT 15 a 9 c 4 g 11 t
ORIGIN

Query Match 65.0%; Score 13; DB 189; Length 39;
 Best Local Similarity 100.0%; Pred. No. 2.1e+04;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 tgcttaaaagctt 18
 |||||
Db 14 TGCCTTAAGCTT 2

RESULT 4
LOCUS TA37H11P 37 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 37h11, forward sequence,
 genomic survey sequence.
ACCESSION AL453069
VERSION AL453069.1 GI:11854580
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
 Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

REFERENCE 1 (bases 1 to 37)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
 Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
 Melville,S.E., Rajandream,M.A. and Barrell,B.G.
 Direct Submission
 Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
 project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
 Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
 nh@sanger.ac.uk

COMMENT Constructed at the Institute for Genomic Research (TIGR),
 Rockville, MD. Genomic DNA isolated from a cloned population of
 trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
 to give a tight size distribution (4 kb). The v + i method used for the library construction is
 described in detail in Smith, H. and Venter, J.C. (Making small
 insert libraries for whole genome shotgun sequencing projects. In
 Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Barrell, Oxford University Press, 1999).
 Email: nelsayed@tigr.org

Details of T. brucei sequencing at the Sanger Centre are available
 at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES Location/Qualifiers
 1..37
source

/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="37h11"

BASE COUNT 8 a 6 c 6 g 17 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 258; Length 37;
Best Local Similarity 87.5%; Pred. No. 2.7e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 tgatgcttaaaagctt 18
||||| || |||||
Db 3 TGATGCTTGAATGCTT 18

RESULT 5

AZ833603 30 bp DNA GSS 20-FEB-2001
LOCUS 2M0115J16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0115J16 R, DNA sequence.

ACCESSION AZ833603
VERSION AZ833603.1 GI:13003511
KEYWORDS GSS.

SOURCE house musculus
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 30)

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL Plasmid inserts
COMMENT Unpublished (2000)

CONTACT: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0115 row: J. column: 16
Seq primer: CACACGGAACACGCTATGACC
Class: plasmid ends

High quality sequence stop: 30.
Location/Qualifiers

FEATURES

source 1. .30
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0115J16"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent *E. coli* XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 11 a 6 c 4 g 9 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 250; Length 30;
Best Local Similarity 82.4%; Pred. No. 5.2e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 catgatgcttaaaagct 17
| | | | | | | | | | | | | | |
Db 5 CCTTATACTTAAAGCT 21

RESULT 6

TA52H11Q 31 bp DNA GSS 13-DEC-2000
LOCUS T. brucei sheared genomic DNA clone 52h11, reverse sequence,
DEFINITION genomic survey sequence.

ACCESSION AL456059
VERSION AL456059.1 GI:11856930
KEYWORDS GSS.

SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei

REFERENCE Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.

AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,
Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,
Melville, S.E., Rajandream, M.A. and Barrell, B.G.
Direct Submission

TITLE Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
JOURNAL project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
nh@sanger.ac.uk

COMMENT Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source 1. .31
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="52h11"

BASE COUNT 9 a 5 c 6 g 11 t
ORIGIN

Query Match 61.0%; Score 12.2; DB 258; Length 31;
Best Local Similarity 82.4%; Pred. No. 5.2e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 tgatgcttaaaagctta 19
| | | | | | | | | | | | | | |
Db 6 TCAGGGTTAATAGCTTA 22

RESULT 7

AI024143 37 bp mRNA EST 27-AUG-1998
LOCUS AI024143

```

DEFINITION  ov73a08.sl Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1642934
              3' similar to gb:D13748 EUKARYOTIC INITIATION FACTOR 4A-I (HUMAN);,
              mRNA sequence.
ACCESSION   AI024143
VERSION     AI024143.1  GI:3239187
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 37)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE       Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
            , Ph.D.
            DNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www.bio.llnl.gov/dbp/image/image.html

Trace considered overall poor quality
Insert Length: 669 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
    1..37
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1642934"
    /clone_lib="Soares_testis_NHT"
    /sex="male"
    /lab_host="DH10B"
    /note="Vector: p7T73D-Pac (Pharmacia) with a modified
    polylinker. Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
    was prepared from mRNA obtained from Clontech Laboratories
    , Inc., and primed with a Not I - oligo(dT) primer [5'.
    TGTTACCAATCGAAGTGGAGCGCCGCCAATTTTTTTTTTTT 3'].
    Double-stranded cDNA was ligated to Eco RI adaptors
    (Pharmacia), digested with Not I and cloned into the Not I
    and Eco RI sites of the modified pT73 vector. Library
    went through one round of normalization to Cot5 and was
    constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT  12 a 8 c 9 g 8 t
ORIGIN
Query Match      61.0%; Score 12.2; DB 15; Length 37;
Best Local Similarity 82.4%; Pred No. 5.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 catgatgtctaaagact 17
    ||||| ||||| |||
Db 19 CATGATACTTAGAAACT 35

RESULT 8
BF144195 38 bp mRNA EST 24-OCT-2000
LOCUS 601786804F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4014744 5',
DEFINITION mRNA sequence.
ACCESSION BF144195
VERSION BF144195
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 38)

```

```

AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9760 row: k column: 10
            High quality sequence stop: 48.
            Location/Qualifiers
    1..48
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:3923841"

BASE COUNT  9 a 9 c 11 g 9 t
ORIGIN
Query Match      61.0%; Score 12.2; DB 145; Length 38;
Best Local Similarity 82.4%; Pred No. 5.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gatgcttaaaagcttac 20
    ||||| ||||| |
Db 18 GATGCATTAAAGCTTGC 34

RESULT 9
BE896146 48 bp mRNA EST 20-OCT-2000
LOCUS 601438738F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3923841 5',
DEFINITION mRNA sequence.
ACCESSION BE896146
VERSION BE896146.1 GI:10360075
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 48)
            NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTF
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9760 row: k column: 10
            High quality sequence stop: 48.
            Location/Qualifiers
    1..48
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:3923841"

```



```

(clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
11 a 14 c 14 g 9 t

BASE COUNT
ORIGIN

Query Match 61.0%; Score 12.2; DB 141; Length 48;
Best Local Similarity 82.4%; Pred. No. 5.5e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 catgatgcttaaaagct 17
||||| || |||||
Db 19 CATGATCCTCACAGCT 35

RESULT 10
H40544
LOCUS 37 bp mRNA EST 31-JUL-1995
yn87b08_r1 Soares adult brain N2B5HB55Y Homo sapiens cDNA clone
IMAGE:175455 5' similar to gb:X63526 ELONGATION FACTOR I-GAMMA
(HUMAN);, mRNA sequence.
H40544
ACCESSION H40544.1 GI:916596
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 37)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1280
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 1280 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
Location/Qualifiers
1..37
/organism="Homo sapiens"
/db_xref="GDB:3837490"
/db_xref="taxon:9606"
/clone="IMAGE:175455"
/clone_lib="Soares adult brain N2B5HB55Y"
/sex="Male"
/dev_stage="55-year old"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: brain; Vector: pT7R3D (Pharmacia) with a
modified polylinker; Site: 1: Not I; Site: 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TCTTACCATCTGAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7R3 vector

```

Db 20 GATGCTTAACAGCCT 34

RESULT 12
TA213B010/C

LOCUS TA213B010 49 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 213B01, reverse sequence, genomic survey sequence.

ACCESSION AL478475
VERSION AL478475.1 GI:11844434
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

REFERENCE 1 (bases 1 to 49)
AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrel@sanger.ac.uk and nh@sanger.ac.uk

COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsavedetigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/projects/T_brucei/.

FEATURES
source
1..49
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="213B01"

BASE COUNT 17 a 12 c 3 g 17 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 258; Length 49;
Best Local Similarity 86.7%; Pred. No. 8.7e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 atgatgcttaaaagc 16
||||| ||||| ||
Db 19 ATGATGATTAAGGC 5

RESULT 13
T73552

LOCUS T73552 50 bp mRNA EST 02-MAR-1995
DEFINITION Yc35h10.r1 Stragene liver (#937224) Homo sapiens cDNA clone IMAGE:82723 5' similar to gb:J04449 CYTOCHROME P450 IIIA4 (HUMAN);, mRNA sequence.

ACCESSION T73552
VERSION T73552.1 GI:690227
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 50)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B., Chissole, S., Dietrich, N., DuBuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, W., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,

Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)

97044478

CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 873
High quality sequence starts: 1 High quality sequence stops: 1
Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 873 Std Error: 0.00
Seq Primer: M13RP1
High quality sequence stop: 1.
Location/Qualifiers
1..50
/organism="Homo sapiens"
/db_xref="GDB:499780"
/db_xref="taxon:9606"
/clone="IMAGE:82723"
/clone.lib="Stratagene liver (#937224)"
/sex="male"
/dev_stage="49 years old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: liver; Vector: pBluescript SK; Site_1: EcoRI ; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Hepatectomy from normal male caucasian. Average insert size: 1.1 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"

BASE COUNT 10 a 7 c 9 g 17 t
ORIGIN

Query Match 59.0%; Score 11.8; DB 189; Length 50;
Best Local Similarity 72.2%; Pred. No. 8.7e+04;
Matches 13; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 catgatgcttaaaagcctt 18
||||| ||||| |||||
Db 14 CANGANGCTTTAAGATT 31

RESULT 14
TA318D05P/C

LOCUS TA318D05P 32 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 318D05, forward sequence, genomic survey sequence.

ACCESSION AL492652
VERSION AL492652.1 GI:11867479
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

REFERENCE 1 (bases 1 to 32)
AUTHORS Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R., Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L., Melville, S.E., Rajandream, M.A. and Barrell, B.G.
TITLE Direct Submission
JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrel@sanger.ac.uk and nh@sanger.ac.uk

COMMENT Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (

4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source
1..32
/organism="Trypanosoma brucei"
/strain="PREU927"
/db_xref="taxon:5691"
/clone="318d05"
8 a 6 c 5 g 13 t

BASE COUNT
ORIGIN

Query Match 58.0%; Score 11.6; DB 258; Length 32;
Best Local Similarity 77.8%; Pred. No. 1e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 atgatgcttaaaagctta 19
||||| |||||||||
Db 26 ATGAAGTGTAAAGCTTA 9

RESULT 15

AZ487122/c
LOCUS AZ487122 38 bp DNA GSS 05-OCT-2000
DEFINITION IM0316P22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0316P22 F, DNA sequence.

ACCESSION AZ487122
VERSION AZ487122.1 GI:10654557
KEYWORDS GSS.

SOURCE

house mouse
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 38)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0316 row: P column: 22

Seq primer: CGTGTAAACGAGGCCAGT

Class: plasmid ends

High quality sequence stop: 38.

Location/Qualifiers

FEATURES

source
1..38
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0316P22"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 13 a 4 c 9 g 12 t
ORIGIN

Query Match 58.0%; Score 11.6; DB 244; Length 38;
Best Local Similarity 77.8%; Pred. No. 1.1e+05;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 tgatgcttaaaagcttac 20

| |||| |||| |||| ||

Db 30 TAATGCATAAATGCTCAC 13

Search completed: October 2, 2001, 15:01:10
Job time: 10839 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:52 ; Search time 3339.34 Seconds
(without alignments)
69.480 Million cell updates/sec

Title: US-09-757-100B-23
Perfect score: 15
Sequence: 1 gcgggctcacagtgg 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_cm: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_cm: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_vi: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_vil: *
59: gb_vl2: *
60: gb_vl1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_rol: *
95: gb_ro2: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	11.4	76.0	18	9 AR078873	AR078873 Sequence
C 2	11.4	76.0	22	10 I76917	I76917 Sequence 25
C 3	11	73.3	21	10 E26933	E26933 Vascular en
C 4	11	73.3	25	9 AR090736	AR090736 Sequence
C 5	11	73.3	31	10 I11623	I11623 Sequence 8
C 6	11	73.3	37	95 S45430	S45430 V alpha J a
7	10.8	72.0	18	9 AR080704	AR080704 Sequence
8	10.8	72.0	24	9 AR019663	AR019663 Sequence

c 9 10.8 72.0 24 9 AR070473 Sequence
c 10 10.8 72.0 24 9 AR099765 Sequence
c 11 10.8 72.0 24 10 I47711 Sequence 10
c 12 10.8 72.0 29 9 AR082206 Sequence 50
c 13 10.8 72.0 29 10 I78252 Sequence 50
c 14 10.8 72.0 30 10 AX092192 Sequence
c 15 10.8 72.0 36 9 AX007130 Sequence
c 16 10.8 72.0 36 10 E51076 Novel Esche
c 17 10.8 72.0 40 9 AR095613 Sequence
c 18 10.8 72.0 40 9 AR095615 Sequence
c 19 10.4 69.3 21 9 AR049682 Sequence
c 20 10.4 69.3 21 9 AR098915 Sequence
c 21 10.4 69.3 21 10 I79755 Sequence 51
c 22 10.4 69.3 23 9 A60222 Sequence 18
c 23 10.4 69.3 26 9 AR090849 Sequence 18
c 24 10.4 69.3 39 10 I06514 Sequence 18
c 25 10.4 69.3 45 9 MWTGPAIC7
c 26 10.2 68.0 17 10 I46492 Sequence 47
c 27 10.2 68.0 20 9 AR073339 Sequence 6
c 28 10.2 68.0 21 9 A69911 Sequence 6
c 29 10.2 68.0 22 9 AX041988 Sequence
c 30 10.2 68.0 24 9 AR017555 Sequence
c 31 10.2 68.0 24 9 AR084745 Sequence
c 32 10.2 68.0 24 10 I21733 Sequence 8
c 33 10.2 68.0 25 9 A58772 Sequence 18
c 34 10.2 68.0 25 9 AX019995 Sequence
c 35 10.2 68.0 25 9 AX020004 Sequence
c 36 10.2 68.0 25 10 I19529 Sequence 6
c 37 10.2 68.0 25 10 I19530 Sequence 7
c 38 10.2 68.0 25 10 I19532 Sequence 9
c 39 10.2 68.0 25 10 I86209 Sequence 6
c 40 10.2 68.0 25 10 I86210 Sequence 7
c 41 10.2 68.0 25 10 I86212 Sequence 9
c 42 10.2 68.0 26 9 A58770 Sequence 16
c 43 10.2 68.0 26 9 AR017553 Sequence 11
c 44 10.2 68.0 26 10 I19534 Sequence 11
c 45 10.2 68.0 26 10 I21731 Sequence 6

ALIGNMENTS

RESULT 1
AR078873/c
LOCUS AR078873 18 bp DNA PAT 31-AUG-2000
DEFINITION Sequence 17 from patent US 5965370.
ACCESSION AR078873
VERSION AR078873.1 GI:10005619
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 18)
AUTHORS Cowsett,L.M.
TITLE Antisense modulation of RhoG expression
JOURNAL Patent: US 5965370-A 17 12-OCT-1999;
FEATURES
source 1..18
BASE COUNT 1 a 6 c 6 g 5 t
ORIGIN
Query Match 76.0%; Score 11.4; DB 9; Length 18;
Best Local Similarity 92.3%; Pred. No. 4.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 cgggctcacagt 14
||||| |||||
Db 17 CGGGCGCACAGT 5
RESULT 2
Qy 3 gggctcacagt 15
||||| |||||
Db 13 GGGCACACAGT 1
Query Match 76.0%; Score 11.4; DB 10; Length 22;
Best Local Similarity 92.3%; Pred. No. 4.3e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 3 gggctcacagt 15
||||| |||||
Db 13 GGGCACACAGT 1
RESULT 3
E26933/c
LOCUS E26933 21 bp DNA PAT 07-FEB-2001
DEFINITION Vascular endothelial cell growth factor.
ACCESSION E26933
VERSION E26933.1 GI:13026353
KEYWORDS JP 1999169183-A/13.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
1 (bases 1 to 21)
AUTHORS Hiroshi,M.S.C.C., Takashi,S.R.W.W. and Suzuki.
TITLE Vascular endothelial cell growth factor
JOURNAL Patent: JP 1999169183-A 13 29-JUN-1999;
COMMENT AGENCY OF IND SCIENCE & TECHNOL,TOAGOSEI CHEM IND CO LTD
OS Unidentified
PN JP 1999169183-A/13
PD 29-JUN-1999
PF 11-DEC-1999 JP 1997362118
PR
PI HIROSHI MITSUI,SUNIL C KAURU,TAKASHI SUGIHARA,RENU WADAWA, PI
HIDEO SUZUKI
PC C12N15/09,A61K38/22,A61K38/22,A61K38/22,A61K38/22,A61K38/22,
PC C07K14/52,
PC C12N1/21,C12N5/10,C12P21/02/(C12N1/21,C12R1:19),(C12N5/10, PC
C12R1:91),
PC (C12P21/02,C12R1:19),(C12P21/02,C12R1:91),C12N15/00,A61K37/24,
PC A61K37/24,
PC A61K37/24,A61K37/24,A61K37/24,C12N5/00,(C12N5/00,C12R1:91) CC
Strandedness: Single;
CC Topology: Linear;
FH Key 1..21
FT source 1..21
Location/Qualifiers
/organism='Unidentified'.
FEATURES
source 1..21
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 4 a 5 c 6 g 6 t
ORIGIN

Query Match

73.3%; Score 11; DB 10; Length 21;

```

Best Local Similarity 100.0%; Pred. No. 7.8e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggtccacagt 14
Db 11 GGCTCACAGT 1

RESULT 4
AR090736/c
LOCUS AR090736 25 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 856 from patent US 5994076.
ACCESSION AR090736
VERSION AR090736.1 GI:10017491
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS Chenchik,A., Jekhadze,G. and Bibilashvilli,R.
TITLE Methods of assaying differential expression
JOURNAL Patent: US 5994076-A 856 30-NOV-1999;
FEATURES
source Location/Qualifiers
1..25
6 a 9 c 6 g 4 t
BASE COUNT
ORIGIN

Query Match 73.3%; Score 11; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 ggtccacagt 14
Db 11 GGCTCACAGT 1

RESULT 5
I11623/c
LOCUS I11623 31 bp DNA PAT 26-JUL-1995
DEFINITION Sequence 8 from Patent US 5408038.
ACCESSION I11623
VERSION I11623.1 GI:909141
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS Smith,R.S., Curtis,L.K., Koduri,K.R., Witztum,J.L. and Young,S.G.
TITLE Nonnatural apolipoprotein B-100 peptides and apolipoprotein
B-100-apolipoprotein A-I fusion peptides
JOURNAL Patent: US 5408038-A 8 18-APR-1995;
FEATURES
source Location/Qualifiers
1..31
10 a 8 c 7 g 6 t
BASE COUNT
ORIGIN

Query Match 73.3%; Score 11; DB 10; Length 31;
Best Local Similarity 100.0%; Pred. No. 7.3e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ggtccacagt 15
Db 19 GGTCCACAGT 9

RESULT 6
S45430/c
LOCUS S45430 37 bp DNA ROD 31-MAR-1995
DEFINITION V alpha J alpha -T-cell receptor alpha chain [mice, BALB C, adult

day-37 thymus, Genomic, 37 nt].
S45430
S45430.1 GI:1679985
Mus sp. BALB C adult day-37 thymus.
Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 37)
Roth,M.E., Holman,P.O. and Kranz,D.M.
Nonrandom use of J alpha gene segments. Influence of V alpha and J
alpha gene location
J. Immunol. 147 (3), 1075-1081 (1991)
91318150
GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 45430] from the original journal article.
This sequence comes from Figure 2 B.
On Nov 21, 1996 this sequence version replaced gi:1619731.
FEATURES
source Location/Qualifiers
1..37
/organism="Mus sp."
/db_xref="taxon:10095"
1..37
/partial
/gene="VJaag;JJaag;"
/note="T-cell receptor alpha chain"
6 a 10 c 9 g 12 t
BASE COUNT
ORIGIN

Query Match 73.3%; Score 11; DB 95; Length 37;
Best Local Similarity 100.0%; Pred. No. 7.1e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 cggggtcacagt 12
Db 21 CGGGCTCACAG 11

RESULT 7
AR080704
LOCUS AR080704 18 bp DNA PAT 31-AUG-2000
DEFINITION Sequence 9 from patent US 5968826.
ACCESSION AR080704
VERSION AR080704.1 GI:10007434
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS Bennett,C.Frank, Condon,T.P. and Cowser,L.M.
TITLE Antisense inhibition of integrin.alpha.4 expression
JOURNAL Patent: US 5968826-A 9 19-OCT-1999;
FEATURES
source Location/Qualifiers
1..18
2 a 7 c 5 g 4 t
BASE COUNT
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 18;
Best Local Similarity 85.7%; Pred. No. 1.1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gcggggtcacagt 14
Db 1 GCGGCTCTCAGT 14

RESULT 8
AR019663
LOCUS AR019663 24 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 2 from patent US 5783683.

```

```
ACCESSION AR019663
VERSION AR019663.1 GI:3974777
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 24)
AUTHORS Morrison,R.S.
TITLE Antisense oligonucleotides which reduce expression of the FGFR1
JOURNAL Patent: US 5783683-A 2 21-JUL-1998;
FEATURES Location/Qualifiers
source 1..24
BASE COUNT 5 a 6 c 7 g 6 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 24;
Best Local Similarity 85.7%; Pred. NO. 1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 15
||| ||||| |||||
Db 1 CGAGCTCACTGTGG 14

RESULT 9
AR070473/c AR070473 24 bp DNA PAT 18-FEB-2000
LOCUS AR070473
DEFINITION Sequence 9 from patent US 5905185.
ACCESSION AR070473
VERSION AR070473.1 GI:7221361
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 24)
AUTHORS Garner,I., Cottingham,I.R., Temperley,S.M., Foster,D.C.,
Sprecher,C.A. and Prunkard,D.E.
TITLE Protein C production in non-human transgenic mammals
JOURNAL Patent: US 5905185-A 9 18-MAY-1999;
FEATURES Location/Qualifiers
source 1..24
BASE COUNT 5 a 8 c 6 g 5 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 24;
Best Local Similarity 85.7%; Pred. NO. 1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 15
||| ||||| |||||
Db 24 CTGGCTCACAGG 11

RESULT 10
AR099765/c AR099765 24 bp DNA PAT 14-FEB-2001
LOCUS AR099765
DEFINITION Sequence 9 from patent US 6077990.
ACCESSION AR099765
VERSION AR099765.1 GI:12809531
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 24)
AUTHORS Leung,W., Darrow,A. and Andrade-Gordon,P.
TITLE PAR2 modified transgenic mice
JOURNAL Patent: US 6077990-A 9 20-JUN-2000;
FEATURES Location/Qualifiers

source 1..24
BASE COUNT 7 a 7 c 9 g 6 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 29;
Best Local Similarity 85.7%; Pred. NO. 9.9e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gcgggctcacagtgg 14
||| ||||| ||||| |||||
```

```
source 1..24
/organism="unknown"
BASE COUNT 4 a 6 c 8 g 6 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 24;
Best Local Similarity 85.7%; Pred. NO. 1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 15
||| ||||| |||||
Db 15 CCGGCTAACAGTGG 2

RESULT 11
I47711/c I47711 24 bp DNA PAT 07-OCT-1997
LOCUS I47711
DEFINITION Sequence 10 from patent US 5639940.
ACCESSION I47711
VERSION I47711.1 GI:2471676
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 24)
AUTHORS Garner,I., Dalrymple,M.L., Prunkard,D.E. and Foster,D.C.
TITLE Production of fibrinogen in transgenic animals
JOURNAL Patent: US 5639940-A 10 17-JUN-1997;
FEATURES Location/Qualifiers
source 1..24
BASE COUNT 5 a 8 c 6 g 5 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 10; Length 24;
Best Local Similarity 85.7%; Pred. NO. 1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 15
||| ||||| ||||| |||||
Db 24 CTGGCTCACAGG 11

RESULT 12
AR082206/c AR082206 29 bp DNA PAT 31-AUG-2000
LOCUS AR082206
DEFINITION Sequence 50 from patent US 5972704.
ACCESSION AR082206
VERSION AR082206.1 GI:10008932
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 29)
AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and
Thompson,J.D.
TITLE HIV nef targeted ribozymes
JOURNAL Patent: US 5972704-A 50 26-OCT-1999;
FEATURES Location/Qualifiers
source 1..29
BASE COUNT 7 a 7 c 9 g 6 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 29;
Best Local Similarity 85.7%; Pred. NO. 9.9e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gcgggctcacagtgg 14
||| ||||| ||||| |||||
```


Db 15 GCAGGCTCACAGG 2

RESULT 13

LOCUS I78252/c 178252 29 bp DNA PAT 03-APR-1998
DEFINITION Sequence 50 from patent US 5693535.
ACCESSION I78252
VERSION I78252.1 GI:3014406
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 29)
AUTHORS Draper,K.G., Chowrira,B., McSwiggen,J., Stinchcomb,D.T. and Thompson,J.D.
TITLE HIV targeted ribozymes
JOURNAL Patent: US 5693535-A 50 02-DEC-1997;
FEATURES
Location/Qualifiers
source 1..29
BASE COUNT 7 a 7 c 9 g 6 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 10; Length 29;
Best Local Similarity 85.7%; Pred. No. 9.9e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gcgggctcacagt 14

|||||

Db 15 GCAGGCTCACAGG 2

RESULT 14

LOCUS AX092192 30 bp DNA PAT 21-MAR-2001
DEFINITION Sequence 8 from Patent WO0116321.
ACCESSION AX092192
VERSION AX092192.1 GI:13444410
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 30)
AUTHORS Tandon,N., Sun,B., Nakamura,T. and Yamamoto,N.
TITLE Platelet membrane glycoprotein vi (gpvi) dna and protein sequences, and uses thereof
JOURNAL Patent: WO 0116321-A 8 08-MAR-2001;
OTSUKA PHARMACEUTICAL CO., LTD. (JP)
FEATURES
Location/Qualifiers
source 1..30
BASE COUNT 6 a 11 c 10 g 3 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 10; Length 30;
Best Local Similarity 85.7%; Pred. No. 9.8e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 15

|||||

Db 12 CGGCTCAGAGTGG 25

RESULT 15

LOCUS AX007130 36 bp DNA PAT 06-SEP-2000
DEFINITION Sequence 16 from Patent EP0972838.
ACCESSION AX007130

VERSION AX007130.1 GI:9995021
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 36)
AUTHORS Schantz,C. and Kopetzki,E.
TITLE Escherichia coli host/vector system based on antibiotic-free selection by complementation of an auxotrophy
JOURNAL Patent: EP 0972838-A 16 19-JAN-2000;
ROCHE DIAGNOSTICS GMBH (DE)
FEATURES
Location/Qualifiers
source 1..36
BASE COUNT 14 a 7 c 12 g 3 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 36;
Best Local Similarity 85.7%; Pred. No. 9.5e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 15

|||||

Db 12 CGGCTCAAGTGG 25

Search completed: October 2, 2001, 15:56:53
Job time: 14177 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:10 ; Search time 10798.2 Seconds
(without alignments)
13.131 Million cell updates/sec

Title: US-09-757-100B-23
Perfect score: 15
Sequence: 1 gcgggctcacagtgg 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: gb_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estrol1:*
83: em_estrol2:*
84: em_estrol3:*
85: em_estrol4:*
86: em_estrol5:*
87: em_estrol6:*
88: em_estrol7:*
89: em_estrol8:*
90: em_estrol9:*
91: em_estrol10:*
92: em_estrol11:*
93: em_estrol12:*
94: em_estrol13:*
95: em_estrol14:*
96: em_estrol15:*
97: em_estrol16:*
98: em_estrol17:*
99: em_estrol18:*
100: em_estrol19:*
101: em_estrol20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0254 row: A column: 13
 Seq primer: CGTTGTAACACGCGGCAGT
 Class: plasmid ends
 High quality sequence stop: 40.
 Location/Qualifiers
 1. 40

TITLE
 JOURNAL
 COMMENT

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	11.8	78.7	40	244	AZ453043 1M0254A13
c 2	11.8	78.7	50	250	AZ834614 2M0117121
c 3	11.4	76.0	50	107	AU106328
c 4	11	73.3	50	107	AU102352
c 5	11	73.3	50	107	AU102356
c 6	10.8	72.0	27	244	AZ455866 1M0258K21
c 7	10.8	72.0	38	246	AZ591206 1M0401K03
c 8	10.8	72.0	41	12	AA822940 vp30g03.r
c 9	10.8	72.0	44	258	AA822940 T. brucei
c 10	10.8	72.0	48	13	AA934014 om58ell.s
c 11	10.8	72.0	50	107	AU107932
c 12	10.4	69.3	22	258	TA140A04P
c 13	10.4	69.3	25	249	AZ770438 1M0572G03
c 14	10.4	69.3	31	242	AZ345566 1M0080114
c 15	10.4	69.3	35	245	AZ497161 1M0333G22
c 16	10.4	69.3	37	249	AZ786210 2M0031P01
c 17	10.4	69.3	42	242	AZ382803 1M0140B13
c 18	10.2	68.0	28	17	AI216798 qm31h06.x
c 19	10.2	68.0	28	24	AI735009 as44604.x
c 20	10.2	68.0	31	12	AA845818 ak85f01.s
c 21	10.2	68.0	34	250	AZ820869 2M0093E20
c 22	10.2	68.0	35	249	AZ798871 2M0055H22
c 23	10.2	68.0	37	1	AA003068 mg48e08.r
c 24	10.2	68.0	37	23	AI684291 tw57b09.x
c 25	10.2	68.0	40	250	AZ812781 2M0079B17
c 26	10.2	68.0	40	240	AI473661 tm02g06.x
c 27	10.2	68.0	40	159	H99190 yx15d07.s1
c 28	10.2	68.0	46	19	AI376441 tc36b06.x
c 29	10.2	68.0	46	241	AZ307757 1M0010E04
c 30	10.2	68.0	49	244	AZ453639 1M0255J08
c 31	10.2	68.0	50	107	AU102588
c 32	10.2	68.0	50	107	AU102590
c 33	10.2	68.0	50	107	AU103607
c 34	10.2	68.0	50	107	AU103616
c 35	10.2	68.0	50	107	AU105119
c 36	10.2	68.0	50	107	AU105138
c 37	10.2	68.0	50	107	AU105362
c 38	10.2	68.0	50	107	AU105659
c 39	10.2	68.0	50	107	AU106912
c 40	10.2	68.0	50	107	AU106912
c 41	10.6	66.7	21	250	AZ812038 2M0078022
c 42	10.6	66.7	22	243	AZ396435
c 43	10.6	66.7	29	249	AZ766277 1M0161C03
c 44	10.6	66.7	38	1	AA035969 ml70h01.r
c 45	10.6	66.7	44	166	BE311212 601057922
c 46	10.6	66.7	46	144	BF123084 601761846

ALIGNMENTS

RESULT 1
 AZ453043/c
 LOCUS 40 bp DNA GSS 04-OCT-2000
 DEFINITION 1M0254A13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0254A13 F, DNA sequence.
 ACCESSION AZ453043
 VERSION AZ453043.1 GI:10610442
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 40)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
 and Wright,D., Weiss,R.

QY 1 gcgggctcacagtgg 15
 | | | | | | | | | |
 Db 25 GGGGCTCACAGTGG 11

RESULT 2
 AZ834614/c
 LOCUS 50 bp DNA GSS 20-FEB-2001
 DEFINITION 2M0117121F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC2M0117121 F, DNA sequence.

ACCESSION AZ834614
 VERSION AZ834614.1 GI:13004522
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 50)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
 and Wright,D., Weiss,R.

Query Match 78.7%; Score 11.8; DB 244; Length 40;
 Best Local Similarity 86.7%; Pred. No. 1.9e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

BASE COUNT 12 a 11 c 9 g 8 t

ORIGIN
 The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gii4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

TITLE and Wright,D.,Weiss,R.
JOURNAL Mouse whole genome scaffolding with paired end reads from 10kb
COMMENT plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0117 row: I column: 21
 Seq primer: CGTTGTAACGACGCGCCAGT
 Class: plasmid ends
 High quality sequence stop: 50.
FEATURES Location/Qualifiers
 source
 1..50
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG240117121"
 /clone_lib="Mouse 10kb plasmid UUCG1m library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g1147321141gb1AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 14 a 10 c 23 g 3 t
ORIGIN

Query Match 78.7%; Score 11.8; DB 250; Length 50;
 Best Local Similarity 86.7%; Pred. NO. 2e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gggggtcacagtgg 15
 |||||
 Db 28 GCGGGCTCACTCTGG 14

RESULT 3
LOCUS AU106328 50 bp mRNA EST 05-APR-2001
DEFINITION AU106328 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone KAT00883, mRNA sequence.
ACCESSION AU106328
VERSION AU106328.1 GI:13555849
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 50)
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo

TITLE K., Suyama,A. and Sugano,S.
JOURNAL Fine Structural analysis of transcription start sites of human
COMMENT mRNAs using full-length enriched and 5'-end enriched cDNA libraries
 Unpublished (2001)
 Contact: Futaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano ,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
 source
 1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="KAT00883"
 /clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 6 a 12 c 22 g 10 t
ORIGIN

Query Match 76.0%; Score 11.4; DB 107; Length 50;
 Best Local Similarity 92.3%; Pred. NO. 3.3e+04;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ggggtcacagtgg 15
 |||||
 Db 22 GGGCCACAGTGG 34

RESULT 4
LOCUS AU102352 50 bp mRNA EST 05-APR-2001
DEFINITION AU102352 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HRC01304, mRNA sequence.
ACCESSION AU102352
VERSION AU102352.1 GI:13551872
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 50)
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo ,K., Suyama,A. and Sugano,S.
 Fine Structural analysis of transcription start sites of human mRNAs using full-length enriched and 5'-end enriched cDNA libraries
 Unpublished (2001)
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano ,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers
 source
 1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HRC01304"
 /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 7 a 16 c 20 g 7 t
ORIGIN

Query Match 73.3%; Score 11; DB 107; Length 50;
 Best Local Similarity 100.0%; Pred. NO. 5.6e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gcggggtcacaca 11

```

|||||
Db 19 GCGGGCTCACCA 9

RESULT 5
AU102356/c
LOCUS AU102356 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone 05-APR-2001
DEFINITION AD5H00944, mRNA sequence.
ACCESSION AU102356
VERSION AU102356.1 GI:13551876
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Tsunoda,T., Talra,H., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
,K., Suyama,A. and Sugano,S.
TITLE Fine Structural analysis of transcription start sites of human
mRNAs using full-length enriched and 5'-end enriched cDNA libraries
JOURNAL Unpublished (2001)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
source 1..50
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ADSH00944"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 8 a 16 c 20 g 6 t
ORIGIN

Query Match 73.3%; Score 11; DB 107; Length 50;
Best Local Similarity 100.0%; Pred. No. 5.6e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 qcgggctcacca 11
|||||
Db 19 GCGGGCTCACCA 9

RESULT 6
AZ455866
LOCUS AZ455866 27 bp DNA GSS 04-OCT-2000
DEFINITION 1M0258K21F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0258K21 F, DNA sequence.
ACCESSION AZ455866
VERSION AZ455866.1 GI:10613991
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 27)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center

|||||
Db 19 GCGGGCTCACCA 9

RESULT 7
AZ591206
LOCUS AZ591206 38 bp DNA GSS 13-DEC-2000
DEFINITION 1M0401K03F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0401K03 F, DNA sequence.
ACCESSION AZ591206
VERSION AZ591206.1 GI:11713312
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 38)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0258 row: K column: 21
Seq primer: CGTTGTAACGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 27.
Location/Qualifiers
1..27
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0258K21"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (G14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 6 a 5 c 10 g 6 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 244; Length 27;
Best Local Similarity 85.7%; Pred. No. 6.9e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cgggctcacagtgg 15
|||||
Db 1 CCGGGTCACATGG 14

RESULT 7
AZ591206
LOCUS AZ591206 38 bp DNA GSS 13-DEC-2000
DEFINITION 1M0401K03F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0401K03 F, DNA sequence.
ACCESSION AZ591206
VERSION AZ591206.1 GI:11713312
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 38)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center

```

insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).

Email: nelsayedetigr.org
 Details of T. brucei sequencing at the Sanger Centre are available
 at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source
 1. .44
 /organism="Trypanosoma brucei"
 /strain="TREU927"
 /db_xref="taxon:5691"
 /clone="233a12"
 12 a 10 c 12 g 10 t
 BASE COUNT 12 a 10 c 12 g 10 t
 ORIGIN
 Query Match 72.0%; Score 10.8; DB 258; Length 44;
 Best Local Similarity 85.7%; Pred. No. 7.2e+04;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gcgggtcacagtgg 14
 ||||| ||||| |
 Db 16 GCGGGATCACAGGG 29

RESULT 10
 AA934014/c
 LOCUS
 DEFINITION
 om58e11.s1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1551404 3'
 similar to gb:X01683 ALPHA-1-ANTITRYPSIN PRECURSOR (HUMAN);, mRNA
 sequence.
 AA934014.1 GI:3090282
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 48)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncilogap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40m13 fwd. ET from Amersham.

FEATURES
 source
 1. .48
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1551404"
 /clone_lib="NCI_CGAP_GC4"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; 1st strand cDNA was prepared from 3 pooled
 germ cell tumors, and was then primed with a Not I -
 oligo(dT) primer. Double-stranded cDNA was ligated to Eco
 RI adaptors (Pharmacia), digested with Not I and cloned
 into the Not I and Eco RI sites of the modified pT7T3
 vector. Library is normalized. Library was constructed by
 Bento Soares and M. Fatima Bonaldo."
 7 a 16 c 15 g 10 t
 BASE COUNT 7 a 16 c 15 g 10 t
 ORIGIN

Query Match 72.0%; Score 10.8; DB 13; Length 48;
 Best Local Similarity 85.7%; Pred. No. 7.3e+04;

Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 cggggtcacagtgg 15
 ||||| ||||| |||||
 Db 30 CCGGGTCACAGAGG 17

RESULT 11
 AU107932
 LOCUS
 DEFINITION
 AU107932 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 ZRV62008, mRNA sequence.

ACCESSION AU107932
 VERSION AU107932.1 GI:13557454
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 50)
 AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
 ,K., Suyama,A. and Sugano,S.
 FINE Structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries
 Unpublished (2001)
 CONTACT: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 ,S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
 source
 1. .50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ZRV62008"
 /clone_lib="Sugano Homo sapiens cDNA library"
 6 a 19 c 15 g 9 t 1 others
 BASE COUNT 6 a 19 c 15 g 9 t 1 others
 ORIGIN

Query Match 72.0%; Score 10.8; DB 107; Length 50;
 Best Local Similarity 85.7%; Pred. No. 7.3e+04;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cggggtcacagtgg 15
 ||||| ||||| |||||
 Db 19 CCGGGCCCGCAGTGG 32

RESULT 12
 TA140A04P/c
 LOCUS
 DEFINITION
 T. brucei sheared genomic DNA clone 140a04, forward sequence,
 genomic survey sequence.

ACCESSION AL466405
 VERSION AL466405.1 GI:11835760
 KEYWORDS GSS.
 SOURCE Trypanosoma brucei.
 ORGANISM Trypanosoma brucei

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma
 1 (bases 1 to 22)
 AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
 Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
 Melville,S.E., Rajandream,M.A. and Barrell,B.G.
 Direct Submission
 Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
 project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
 Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and

nhl@sanger.ac.uk
 Constructed at the Institute for Genomic Research (TIGR),
 Rockville, MD. Genomic DNA isolated from a cloned population of
 Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
 to give a tight size distribution (4 kb). The v + i method used for the library construction is
 described in detail in Smith, H. and Venter, J.C. (Making small
 insert libraries for whole genome shotgun sequencing projects. In
 Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Barrell, Oxford University Press, 1999).
 Email: nelsayed@tigr.org
 Details of T. brucei sequencing at the Sanger Centre are available
 at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source
 1. .22
 /organism="Trypanosoma brucei"
 /strain="TREU927"
 /db_xref="taxon:5691"
 /clone="140a04"
 3 a 6 c 10 g 3 t

BASE COUNT

ORIGIN
 3 a 6 c 10 g 3 t

Query Match 69.3%; Score 10.4; DB 258; Length 22;
 Best Local Similarity 91.7%; Pred. No. 1.1e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ggctcacagtgg 15
 ||| |||||
 Db 13 GCGGCACAGTGG 2

RESULT 13

AZ770438/c
 LOCUS AZ770438 25 bp DNA GSS 16-FEB-2001
 DEFINITION IM0572G03F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 clone UUGCLM0572G03 F, DNA sequence.
 ACCESSION AZ770438
 VERSION AZ770438.1 GI:12891623
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 25)
 REFERENCE
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 , M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0572 row: G column: 03
 Seq primer: CGTTGTAACGACGCGCAGT
 Class: plasmid ends
 High quality sequence stop: 25.

FEATURES

source
 1. .25
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM0572G03"
 /clone_lib="Mouse 10kb plasmid UUGCLM library"
 /sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gil4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 5 a 5 c 9 g 6 t

BASE COUNT

ORIGIN
 5 a 5 c 9 g 6 t

Query Match 69.3%; Score 10.4; DB 249; Length 25;
 Best Local Similarity 91.7%; Pred. No. 1.1e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ggctcacagtgg 15
 | |||||
 Db 15 GTCTCACAGTGG 4

RESULT 14

AZ345566/c
 LOCUS AZ345566 31 bp DNA GSS 29-SEP-2000
 DEFINITION IM0080I14F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 clone UUGCLM0080I14 F, DNA sequence.
 ACCESSION AZ345566
 VERSION AZ345566.1 GI:10424803
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 31)
 REFERENCE
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
 , M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0080 row: I column: 14
 Seq primer: CGTTGTAACGACGCGCAGT
 Class: plasmid ends
 High quality sequence stop: 31.

FEATURES
 source
 1. .31
 Location/Qualifiers
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM0080I14"
 /clone_lib="Mouse 10kb plasmid UUGCLM library"

```

/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      6 a      9 c      6 g      10 t
ORIGIN

```

```

Query Match      69.3%; Score 10.4; DB 242; Length 31;
Best Local Similarity 91.7%; Pred. No. 1.2e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 3 gggtccacagt 14
   ||| |||||
Db 20 GGGATCAGTGG 9

```

RESULT 15 AZ497161/c

```

LOCUS      35 bp      DNA      GSS      05-OCT-2000
DEFINITION 1M0333G22R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0333G22 R, DNA sequence.

```

```

ACCESSION  A2497161
VERSION     A2497161
KEYWORDS   GSS.
SOURCE     house mouse.

```

ORGANISM

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 35)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

```

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

```

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

```

```

Tel: 801 585 5606
Fax: 801 585 7177

```

```

Email: ddunn@genetics.utah.edu

```

```

Insert Length: 10000 Std Error: 0.00

```

```

Plate: 0333 row: G column: 22

```

```

Seq primer: CACACAGGAACACGCTATGACC

```

```

Class: plasmid ends

```

```

High quality sequence stop: 35.

```

```

Location/Qualifiers

```

FEATURES

```

source

```

```

1..35

```

```

/organism="Mus musculus"

```

```

/strain="C57BL/6J"

```

```

/db_xref="taxon:10090"

```

```

/clone="UUGC1M0333G22"

```

```

/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      10 a      9 c      9 g      7 t
ORIGIN

```

```

Query Match      69.3%; Score 10.4; DB 245; Length 35;
Best Local Similarity 91.7%; Pred. No. 1.2e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 4 gggtccacagtgg 15
   ||| |||||
Db 14 GTCTCACAGTGG 3

```

```

Search completed: October 2, 2001, 15:01:11
Job time: 10840 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:53 ; Search time 3339.34 Seconds
(without alignments)
59.480 Million cell updates/sec

Title: US-09-757-100B-30
Perfect score: 15
Sequence: 1 taagcagctgccatt 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_vi: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_vil: *
59: gb_vil2: *
60: gb_vil2: *
61: gb_vil2: *
62: gb_vil2: *
63: gb_vil2: *
64: gb_vil2: *
65: gb_vil2: *
66: gb_vil2: *
67: gb_vil2: *
68: gb_vil2: *
69: gb_vil2: *
70: gb_vil2: *
71: gb_vil2: *
72: gb_vil2: *
73: gb_vil2: *
74: gb_vil2: *
75: gb_vil2: *
76: gb_vil2: *
77: gb_vil2: *
78: gb_vil2: *
79: gb_vil2: *
80: gb_vil2: *
81: gb_vil2: *
82: gb_vil2: *
83: gb_vil2: *
84: gb_vil2: *
85: gb_vil2: *
86: gb_vil2: *
87: gb_vil2: *
88: gb_vil2: *
89: gb_vil2: *
90: gb_vil2: *
91: gb_vil2: *
92: gb_vil2: *
93: gb_vil2: *
94: gb_vil2: *
95: gb_vil2: *
96: gb_vil2: *
97: gb_vil2: *
98: em_ba3: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	13.4	89.3	18	10	I44719 Sequence 45
c 2	13.4	89.3	18	10	I52125 Sequence 45
c 3	11.8	78.7	50	9	AR032874 Sequence
c 4	11.8	78.7	50	9	AR032875 Sequence
c 5	11.8	78.7	50	10	I29614 Sequence 48
c 6	11.8	78.7	50	10	I29615 Sequence 48
c 7	11.8	78.7	50	10	I91288 Sequence 48
c 8	11.8	78.7	50	10	I91289 Sequence 48

```

c 9 11.4 76.0 21 10 AX094845
c 10 11.4 76.0 22 9 A80791
c 11 11.4 76.0 22 9 AX022755
c 12 11.4 76.0 25 9 AX016058
c 13 11.4 76.0 25 9 AR075530
c 14 11.4 76.0 25 9 AX019566
c 15 11.4 76.0 26 10 E04774
c 16 11.4 76.0 30 9 A84595
c 17 11.4 76.0 30 9 AR079892
c 18 11.4 76.0 39 9 AR032341
c 19 11.4 76.0 39 9 AR080603
c 20 11.4 76.0 41 9 A82505
c 21 11.4 76.0 41 9 A82526
c 22 11 73.3 20 45 E10367
c 23 11 73.3 22 9 AR030684
c 24 11 73.3 27 9 A83874
c 25 11 73.3 28 9 AR026168
c 26 11 73.3 28 9 AR026220
c 27 11 73.3 28 9 AR026234
c 28 11 73.3 28 9 AR026248
c 29 11 73.3 28 10 E61327
c 30 11 73.3 28 10 182910
c 31 11 73.3 28 10 182962
c 32 11 73.3 28 10 182976
c 33 11 73.3 28 10 182990
c 34 11 73.3 29 10 146944
c 35 11 73.3 30 9 A12174
c 36 11 73.3 32 10 106269
c 37 11 73.3 35 10 174286
c 38 11 73.3 39 9 A28773
c 39 11 73.3 39 9 A57638
c 40 11 73.3 39 9 AR088301
c 41 11 73.3 41 9 AR032336
c 42 11 73.3 45 9 AX048222
c 43 11 73.3 45 9 AX049383
c 44 11 73.3 45 9 AX052723
c 45 11 73.3 45 9 AX053164

```

ALIGNMENTS

```

RESULT 1
144719/c
LOCUS 144719 18 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 45 from patent US 5635384.
ACCESSION 144719
VERSION 144719.1 GI:2469432
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 18)
AUTHORS Walsh,T.A., Hey,T.D. and Morgan,A.E.R.
TITLE Ribosome-inactivating proteins, inactive precursor forms thereof, a
process for making and a method of using
JOURNAL Patent: US 5635384-A 45 03-JUN-1997;
FEATURES
source
BASE COUNT 5 a 4 c 4 g 5 t
ORIGIN

Query Match 89.3%; Score 13.4; DB 10; Length 18;
Best Local Similarity 93.3%; Pred. No. 1.le+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
| | | | | | | | | |
Db 17 TAAGCAGCTGCAATT 3

RESULT 4
AR032875
LOCUS AR032875 50 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 487 from patent US 5869241.
ACCESSION AR032875
VERSION AR032875.1 GI:5948480
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 50)

```

```

RESULT 2
152125/c
LOCUS 152125 18 bp DNA PAT 07-OCT-1997
DEFINITION Sequence 45 from patent US 5646026.
ACCESSION 152125
VERSION 152125.1 GI:2473326
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 18)
AUTHORS Walsh,T.A., Hey,T.D. and Morgan,A.E.R.
TITLE Ribosome-inactivating proteins, inactive precursor forms thereof, a
process for making and a method of using
JOURNAL Patent: US 5646026-A 45 08-JUL-1997;
FEATURES
source
BASE COUNT 5 a 4 c 4 g 5 t
ORIGIN

Query Match 89.3%; Score 13.4; DB 10; Length 18;
Best Local Similarity 93.3%; Pred. No. 1.le+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
| | | | | | | | | |
Db 17 TAAGCAGCTGCAATT 3

RESULT 3
AR032874
LOCUS AR032874 50 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 486 from patent US 5869241.
ACCESSION AR032874
VERSION AR032874.1 GI:5948479
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 50)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding
molecule
JOURNAL Patent: US 5869241-A 486 09-FEB-1999;
FEATURES
source
BASE COUNT 9 a 14 c 13 g 14 t
ORIGIN

Query Match 78.7%; Score 11.8; DB 9; Length 50;
Best Local Similarity 86.7%; Pred. No. 1.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
| | | | | | | | | |
Db 24 TAAGCAGCTGCTTTT 38

RESULT 4
AR032875
LOCUS AR032875 50 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 487 from patent US 5869241.
ACCESSION AR032875
VERSION AR032875.1 GI:5948480
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 50)

```

AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding molecule

JOURNAL Patent: US 5869241-A 487 09-FEB-1999;

FEATURES Location/Qualifiers
source
1..50

BASE COUNT 8 a 13 c 13 g 16 t

ORIGIN

Query Match 78.7%; Score 11.8; DB 9; Length 50;

Best Local Similarity 86.7%; Pred. No. 1.4e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 15

||||| ||||| ||

Db 24 TAAGCAGCTGCTTTT 38

RESULT 5

I29614

LOCUS I29614 50 bp DNA PAT 06-FEB-1997

DEFINITION Sequence 486 from patent US 5578444.

ACCESSION I29614

VERSION I29614.1 GI:1820405

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 50)

AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.

TITLE Sequence-directed DNA-binding molecules compositions and methods

JOURNAL Patent: US 5578444-A 486 26-NOV-1996;

FEATURES Location/Qualifiers
source
1..50

BASE COUNT 9 a 14 c 13 g 14 t

ORIGIN

Query Match 78.7%; Score 11.8; DB 10; Length 50;

Best Local Similarity 86.7%; Pred. No. 1.4e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 15

||||| ||||| ||

Db 24 TAAGCAGCTGCTTTT 38

RESULT 6

I29615

LOCUS I29615 50 bp DNA PAT 06-FEB-1997

DEFINITION Sequence 487 from patent US 5578444.

ACCESSION I29615

VERSION I29615.1 GI:1820406

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 50)

AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.

TITLE Sequence-directed DNA-binding molecules compositions and methods

JOURNAL Patent: US 5578444-A 487 26-NOV-1996;

FEATURES Location/Qualifiers
source
1..50

BASE COUNT 8 a 13 c 13 g 16 t

ORIGIN

Query Match 78.7%; Score 11.8; DB 10; Length 50;

Best Local Similarity 86.7%; Pred. No. 1.4e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 15

||||| ||||| ||

Db 24 TAAGCAGCTGCTTTT 38

RESULT 7

I91288

LOCUS I91288 50 bp DNA PAT 01-DEC-1998

DEFINITION Sequence 486 from patent US 5726014.

ACCESSION I91288

VERSION I91288.1 GI:3935758

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 50)

AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.

TITLE Screening assay for the detection of DNA-binding molecules

JOURNAL Patent: US 5726014-A 486 10-MAR-1998;

FEATURES Location/Qualifiers
source
1..50

BASE COUNT 9 a 14 c 13 g 14 t

ORIGIN

Query Match 78.7%; Score 11.8; DB 10; Length 50;

Best Local Similarity 86.7%; Pred. No. 1.4e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 15

||||| ||||| ||

Db 24 TAAGCAGCTGCTTTT 38

RESULT 8

I91289

LOCUS I91289 50 bp DNA PAT 01-DEC-1998

DEFINITION Sequence 487 from patent US 5726014.

ACCESSION I91289

VERSION I91289.1 GI:3935759

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 50)

AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.

TITLE Screening assay for the detection of DNA-binding molecules

JOURNAL Patent: US 5726014-A 487 10-MAR-1998;

FEATURES Location/Qualifiers
source
1..50

BASE COUNT 8 a 13 c 13 g 16 t

ORIGIN

Query Match 78.7%; Score 11.8; DB 10; Length 50;

Best Local Similarity 86.7%; Pred. No. 1.4e+04;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 15

||||| ||||| ||

Db 24 TAAGCAGCTGCTTTT 38

RESULT 9

AX094845/C

LOCUS AX094845 21 bp DNA PAT 30-MAR-2001

DEFINITION Sequence 23 from Patent W00118250.

ACCESSION AX094845

VERSION AX094845.1 GI:13511048

```

KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            1 (bases 1 to 21)
TITLE       Lander,E.S., Gargill,M., Ireland,J.S., Bolk,S., Daley,G.O. and
JOURNAL     McCarthy,J.J.
            Single nucleotide polymorphisms in genes
FEATURES    Patent: WO 0118250-A 23 15-MAR-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
            Pharmaceuticals, Inc. (US)
            Location/Qualifiers
            1..21
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT  6 a      6 c      4 g      4 t      1 others
ORIGIN      1..21
            6 a      6 c      4 g      4 t      1 others

Query Match      76.0%; Score 11.4; DB 10; Length 21;
Best Local Similarity 80.0%; Pred. No. 2.5e+04;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
    |||:|||||
Db 17 TAGGCARCTGCGATT 3

RESULT 10
LOCUS      A80791      22 bp      DNA      PAT      21-JAN-2000
DEFINITION Sequence 46 from Patent EP0922767.
ACCESSION  A80791
VERSION     A80791.1 GI:6731415
KEYWORDS   .
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 22)
AUTHORS     Tucker,W. and Huttner,E.
TITLE       Ribozymes capable of conferring resistance to potyvirus infection,
            and plants expressing said ribozymes
JOURNAL     Patent: EP 0922767-A 46 16-JUN-1999;
            GENE SHEARS PTY LTD (AU)
FEATURES    Location/Qualifiers
            1..22
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT  5 a      5 c      6 g      6 t
ORIGIN      1..22

Query Match      76.0%; Score 11.4; DB 9; Length 22;
Best Local Similarity 92.3%; Pred. No. 2.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 agcagctgccatt 15
    |||:|||||
Db 13 AGCAGCTGCAATT 1

RESULT 11
LOCUS      AX022755      22 bp      DNA      PAT      07-SEP-2000
DEFINITION Sequence 46 from Patent WO928485.
ACCESSION  AX022755
VERSION     AX022755.1 GI:10046280
KEYWORDS   .
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 22)

KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            1 (bases 1 to 21)
TITLE       Lander,E.S., Gargill,M., Ireland,J.S., Bolk,S., Daley,G.O. and
JOURNAL     McCarthy,J.J.
            Single nucleotide polymorphisms in genes
FEATURES    Patent: WO 0118250-A 23 15-MAR-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
            Pharmaceuticals, Inc. (US)
            Location/Qualifiers
            1..21
            /organism="unidentified"
            /db_xref="taxon:9606"
BASE COUNT  6 a      6 c      4 g      4 t      1 others
ORIGIN      1..21
            6 a      6 c      4 g      4 t      1 others

Query Match      76.0%; Score 11.4; DB 10; Length 21;
Best Local Similarity 80.0%; Pred. No. 2.5e+04;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
    |||:|||||
Db 17 TAGGCARCTGCGATT 3

RESULT 10
LOCUS      A80791      22 bp      DNA      PAT      21-JAN-2000
DEFINITION Sequence 46 from Patent EP0922767.
ACCESSION  A80791
VERSION     A80791.1 GI:6731415
KEYWORDS   .
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 22)
AUTHORS     Tucker,W. and Huttner,E.
TITLE       Ribozymes capable of conferring resistance to potyvirus infection,
            and plants expressing said ribozymes
JOURNAL     Patent: EP 0922767-A 46 16-JUN-1999;
            GENE SHEARS PTY LTD (AU)
FEATURES    Location/Qualifiers
            1..22
            /organism="unidentified"
            /db_xref="taxon:32644"
BASE COUNT  5 a      5 c      6 g      6 t
ORIGIN      1..22

Query Match      76.0%; Score 11.4; DB 9; Length 22;
Best Local Similarity 92.3%; Pred. No. 2.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 agcagctgccatt 15
    |||:|||||
Db 13 AGCAGCTGCAATT 1

RESULT 11
LOCUS      AX022755      22 bp      DNA      PAT      07-SEP-2000
DEFINITION Sequence 46 from Patent WO928485.
ACCESSION  AX022755
VERSION     AX022755.1 GI:10046280
KEYWORDS   .
SOURCE      unidentified.
ORGANISM    unidentified.
REFERENCE   1 (bases 1 to 22)

AUTHORS     Tucker,W., Huttner,E., Ignart,F. and Vermeulen,A.
TITLE       Ribozymes capable of conferring resistance to potyvirus infection,
            and plants expressing said ribozymes
JOURNAL     Patent: WO 928485-A 46 10-JUN-1999;
            GENE SHEARS PTY LTD (AU); TUCKER WILLIAM (AU); HUTTNER ERIC (FR);
            IGNART FREDERIC (FR); VERMEULEN AGNES (FR)
FEATURES    Location/Qualifiers
            1..22
            /organism="unidentified"
            /db_xref="taxon:32644"
            /note="WMPC1S"
BASE COUNT  5 a      5 c      6 g      6 t
ORIGIN      1..22
            5 a      5 c      6 g      6 t

Query Match      76.0%; Score 11.4; DB 9; Length 22;
Best Local Similarity 92.3%; Pred. No. 2.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 agcagctgccatt 15
    |||:|||||
Db 13 AGCAGCTGCAATT 1

RESULT 12
LOCUS      AR016058      25 bp      DNA      PAT      05-DEC-1998
DEFINITION Sequence 26 from patent US 5776679.
ACCESSION  AR016058
VERSION     AR016058.1 GI:3972335
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 25)
AUTHORS     Villeponteau,B., Feng,J., Funk,W. and Andrews,W.H.
TITLE       Assays for the DNA component of human telomerase
JOURNAL     Patent: US 5776679-A 26 07-JUL-1998;
FEATURES    Location/Qualifiers
            1..25
            /organism="unknown"
BASE COUNT  8 a      8 c      6 g      3 t
ORIGIN      1..25

Query Match      76.0%; Score 11.4; DB 9; Length 25;
Best Local Similarity 92.3%; Pred. No. 2.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 agcagctgccatt 15
    |||:|||||
Db 13 AGCAGCTGACATT 25

RESULT 13
LOCUS      AR075530      25 bp      DNA      PAT      30-AUG-2000
DEFINITION Sequence 27 from patent US 5958680.
ACCESSION  AR075530
VERSION     AR075530.1 GI:10002278
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 25)
AUTHORS     Villeponteau,B., Feng,J., Funk,W. and Andrews,W.H.
TITLE       Mammalian telomerase
JOURNAL     Patent: US 5958680-A 27 28-SEP-1999;
FEATURES    Location/Qualifiers
            1..25
            /organism="unknown"
BASE COUNT  8 a      8 c      6 g      3 t

```


ORIGIN

Query Match 76.0%; Score 11.4; DB 9; Length 25;
Best Local Similarity 92.3%; Pred. No. 2.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 agcagctgccatt 15
|||||
Db 13 AGCAGCTGACATT 25

RESULT 14

AX019566
LOCUS AX019566 25 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 20 from Patent WO9938964.
ACCESSION AX019566
VERSION AX019566.1 GI:10043480
SOURCE .
KEYWORDS synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 25)
AUTHORS Keith,W.N.
TITLE Promoter regions of the mouse and human telomerase rna component
genes
JOURNAL Patent: WO 9938964-A 20 05-AUG-1999;
KEITH WILLIAM NICOL (GB); CANCER RES CAMPAIGN TECH (GB)

FEATURES

source
1. .25
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer"

BASE COUNT 6 a 7 c 9 g 3 t

ORIGIN

Query Match 76.0%; Score 11.4; DB 9; Length 25;
Best Local Similarity 92.3%; Pred. No. 2.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 agcagctgccatt 15
|||||
Db 13 AGCAGCTGACATT 25

RESULT 15

E04774/c
LOCUS E04774 26 bp DNA PAT 29-SEP-1997
DEFINITION Synthetic DNA for site directed mutagenesis of protease derived
from Bacillus NKS-21.
ACCESSION E04774
VERSION E04774.1 GI:2172970
KEYWORDS JP 1993091876-A/13.
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 26)
AUTHORS Takinishi,E., Kakinuma,S., Takemoto,A., Miyoda,Y. and Fukuyama,S.
TITLE PROTEASE, ITS PRODUCTION AND USE
JOURNAL Patent: JP 1993091876-A 13 16-APR-1993;
SHOWA DENKO KK

COMMENT
OS Artificial gene
OC Artificial sequence; Genes.
OS Bacillus NKS-21
PN JP 1993091876-A/13
PD 16-APR-1993
PF 02-OCT-1991 JP 1991280313
PI TAKINISHI EIKO, KAKINUMA SHINJI, TAKEMOTO AYANO, PI MIYODA
YOSHIAKI,
PI FUKUYAMA SHIRO
PC C12N9/54,C11D3/386,C12N1/21,C12N15/57,(C12N9/54,C12R1:07), PC
(C12N1/21,

PC C12R1:07);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: NO;
CC anti-sense: NO;
FEATURES
Location/Qualifiers
source
1. .26
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 7 a 5 c 5 g 9 t
ORIGIN

Query Match 76.0%; Score 11.4; DB 10; Length 26;
Best Local Similarity 92.3%; Pred. No. 2.5e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 agcagctgccatt 15
|||||
Db 18 AGCAGCTGCAATT 6

Search completed: October 2, 2001, 15:56:53
Job time: 14177 sec

,
,
,
.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:11 ; Search time 10798.2 Seconds
(without alignments)
13.131 Million cell updates/sec

Title: US-09-757-100B-30

Perfect score: 15

Sequence: 1 taagcagctgccatt 15

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:**

- 1: gb_est1:**
- 2: gb_est2:**
- 3: gb_est3:**
- 4: gb_est4:**
- 5: gb_est5:**
- 6: gb_est6:**
- 7: gb_est7:**
- 8: gb_est8:**
- 9: gb_est9:**
- 10: gb_est10:**
- 11: gb_est11:**
- 12: gb_est12:**
- 13: gb_est13:**
- 14: gb_est14:**
- 15: gb_est15:**
- 16: gb_est16:**
- 17: gb_est17:**
- 18: gb_est18:**
- 19: gb_est19:**
- 20: gb_est20:**
- 21: gb_est21:**
- 22: gb_est22:**
- 23: gb_est23:**
- 24: gb_est24:**
- 25: gb_est33:**
- 26: gb_est34:**
- 27: gb_est35:**
- 28: gb_est36:**
- 29: gb_est37:**
- 30: gb_est38:**
- 31: gb_est39:**
- 32: gb_est40:**
- 33: em_estba:**
- 34: em_estfun:**
- 35: em_esthum1:**
- 36: em_esthum2:**
- 37: em_esthum3:**
- 38: em_esthum4:**
- 39: em_esthum5:**
- 40: em_esthum6:**
- 41: em_esthum7:**
- 42: em_esthum8:**
- 43: em_esthum9:**

- 44: em_esthum10:**
- 45: em_esthum11:**
- 46: em_esthum12:**
- 47: em_esthum13:**
- 48: em_esthum14:**
- 49: em_esthum15:**
- 50: em_esthum16:**
- 51: em_esthum17:**
- 52: em_esthum18:**
- 53: em_esthum19:**
- 54: em_esthum20:**
- 55: em_esthum21:**
- 56: em_esthum22:**
- 57: em_esthum23:**
- 58: em_esthum24:**
- 59: em_esthum25:**
- 60: em_esthum26:**
- 61: em_esthum27:**
- 62: em_esthum28:**
- 63: em_estin1:**
- 64: em_estin2:**
- 65: em_estin3:**
- 66: em_estin4:**
- 67: em_estin5:**
- 68: em_estom1:**
- 69: em_estom2:**
- 70: em_estov1:**
- 71: em_estov2:**
- 72: em_estpl1:**
- 73: em_estpl2:**
- 74: em_estpl3:**
- 75: em_estpl4:**
- 76: em_estpl5:**
- 77: em_estpl6:**
- 78: em_estpl7:**
- 79: em_estpl8:**
- 80: em_estpl9:**
- 81: em_estpl10:**
- 82: em_estro1:**
- 83: em_estro2:**
- 84: em_estro3:**
- 85: em_estro4:**
- 86: em_estro5:**
- 87: em_estro6:**
- 88: em_estro7:**
- 89: em_estro8:**
- 90: em_estro9:**
- 91: em_estro10:**
- 92: em_estro11:**
- 93: em_estro12:**
- 94: em_estro13:**
- 95: em_estro14:**
- 96: em_estro15:**
- 97: em_estro16:**
- 98: em_estro17:**
- 99: em_estro18:**
- 100: em_estro19:**
- 101: em_estro20:**
- 102: gb_est25:**
- 103: gb_est26:**
- 104: gb_est27:**
- 105: gb_est28:**
- 106: gb_est29:**
- 107: gb_est30:**
- 108: gb_est31:**
- 109: gb_est32:**
- 110: gb_est41:**
- 111: gb_est42:**
- 112: gb_est43:**
- 113: gb_est44:**
- 114: gb_est45:**
- 115: gb_est46:**
- 116: gb_est47:**

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

JOURNAL COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 660 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES
source
1. 41
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1420657"
/clone.lib="NCI-CGAP_Col2"
/sex="mixed"
/tissue_type="colon tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: colon; Vector: Bluescript SK-; Site: 1: EcoRI
; Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. Pooled colon tumors. 5' adaptor sequence: 5'
GAATTCGACGACGAG 3' adaptor sequence: 5'-
CTCCAGGTTTTTTTTTTT 3' Average insert size: 1.2 kb."
BASE COUNT 10 a 6 c 11 g 14 t
ORIGIN

Query Match 78.7%; Score 11.8; DB 12; Length 41;
Best Local Similarity 86.7%; Pred. No. 3.2e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
||||||| |||||
Db 19 TAAGCAGCCACCAT 5

RESULT 2
AA662608 50 bp mRNA EST 12-NOV-1997
LOCUS nr16a12.s1 NCI-CGAP_Ew1 Homo sapiens cDNA clone IMAGE:1168126
DEFINITION similar to SW:RL24_HUMAN P38663 60S RIBOSOMAL PROTEIN L24. ;, mRNA
sequence.
ACCESSION AA662608
VERSION AA662608.1 GI:2615457
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,
Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40ml3 fwd. ET from Amersham.
Location/Qualifiers

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	11.8	78.7	41	12	AA829551	of09h01.s
c 2	11.8	78.7	50	10	AA662608	nr16a12.s
c 3	11.4	76.0	47	143	AA794942	vr05g03.r
c 4	11.4	76.0	45	123	BF036368	601460534
c 5	11	73.3	47	158	H59245	yr04b03.s1
c 6	11	73.3	50	107	AU103473	AU103473
c 7	11	73.3	50	107	AU104774	AU104774
c 8	10.8	72.0	19	244	AZ481008	IM0302N15
c 9	10.8	72.0	28	258	TA355F03Q	AL497072 T. brucei
c 10	10.8	72.0	40	12	AA806226	oe29f12.s
c 11	10.8	72.0	41	258	TAJ26G07Q	AL463426 T. brucei
c 12	10.8	72.0	42	113	AW249855	2821534.3
c 13	10.8	72.0	44	9	AA553766	nk29h01.s
c 14	10.8	72.0	44	9	AA600013	ag29h11.s
c 15	10.8	72.0	44	11	AA713885	nv79d06.s
c 16	10.8	72.0	44	14	AA989114	or75g04.s
c 17	10.8	72.0	50	20	AI476044	tl97b11.x
c 18	10.8	72.0	50	107	AU102666	AU102666
c 19	10.8	72.0	50	107	AU103462	AU103462
c 20	10.8	72.0	50	107	AU104066	AU104066
c 21	10.8	72.0	50	107	AU107071	AU107071
c 22	10.4	69.3	20	243	AZ435787	IM0223G02
c 23	10.4	69.3	21	242	AZ374998	IM0128H11
c 24	10.4	69.3	22	243	AZ430117	IM0214M24
c 25	10.4	69.3	22	249	AZ805989	2M0067C13
c 26	10.4	69.3	26	247	AZ632821	IM0487K12
c 27	10.4	69.3	31	15	AI032592	ow73g08.s
c 28	10.4	69.3	31	244	AZ448456	IM0246H08
c 29	10.4	69.3	32	258	TA371C10P	AL495613 T. brucei
c 30	10.4	69.3	34	4	AA282064	AZ282064 zt02e01.s
c 31	10.4	69.3	34	249	AZ785014	2M0028D03
c 32	10.4	69.3	38	249	AZ759713	IM0552E14
c 33	10.4	69.3	41	241	AZ327077	IM0050C16
c 34	10.4	69.3	41	247	AZ634745	IM0490O12
c 35	10.4	69.3	42	249	AZ760286	IM0553M20
c 36	10.4	69.3	45	188	TI7561	mps v160 fh
c 37	10.4	69.3	48	249	AZ787799	2M0034C06
c 38	10.4	69.3	49	1	AA047951	AA047951 mj23h04.r
c 39	10.4	69.3	49	13	AA914273	vy99a03.r
c 40	10.4	69.3	50	107	AU102448	AU102448
c 41	10.4	69.3	50	107	AU103167	AU103167
c 42	10.4	69.3	50	107	AU103390	AU103390
c 43	10.4	69.3	50	107	AU103485	AU103485
c 44	10.4	69.3	50	107	AU103486	AU103486
c 45	10.4	69.3	50	107	AU103487	AU103487

ALIGNMENTS

RESULT 1
AA829551/c
LOCUS of09h01.s1 NCI-CGAP_Col2 Homo sapiens cDNA clone IMAGE:1420657 3'
DEFINITION similar to gb:M33552 LYMPHOCYTE-SPECIFIC PROTEIN LSP1 (HUMAN);,
mRNA sequence.
ACCESSION AA829551
VERSION AA829551.1 GI:2902650
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 41)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index


```

Best Local Similarity 100.0%; Pred. No. 9e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 aagcagctgcc 12
   |||||
Db 36 AAGCAGCTGCC 26

RESULT 8
LOCUS AZ481008 19 bp DNA GSS 04-OCT-2000
DEFINITION IM0302N15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0302N15 R, DNA sequence.
ACCESSION AZ481008
VERSION AZ481008.1 GI:10641989
KEYWORDS GSS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 19)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 309, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0302 row: N column: 15
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 19.
FEATURES
Location/Qualifiers
1..19
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0302N15"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 3 a 4 c 6 g 6 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 244; Length 19;
Best Local Similarity 85.7%; Pred. No. 9.8e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 aagcagctgccatt 15
   |||||
Db 14 AAGCAGCTACCGT 1

RESULT 9
LOCUS TA355F03Q/c 28 bp DNA GSS 13-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 355f03, reverse sequence,
genomic survey sequence.
ACCESSION AL497072
VERSION AL497072.1 GI:11872664
KEYWORDS GSS.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 28)
AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
Melville,S.E., Rajandream,M.A. and Barrell,B.G.
Direct Submission
Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
nh@sanger.ac.uk
COMMENT Constructed at the Institute for Genomic Research (TIGR),
Rockville, MD. Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
to give a tight size distribution (
4 kb). The v + i method used for the library construction is
described in detail in Smith, H. and Venter, J.C. (Making small
insert libraries for whole genome shotgun sequencing projects. In
Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barrell, Oxford University Press, 1999).
Email: nelsayed@tigr.org
Details of T. brucei sequencing at the Sanger Centre are available
at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES
Location/Qualifiers
1..28
/organism="Trypanosoma brucei"
/strain="TREU927"
/db_xref="taxon:5691"
/clone="355f03"
BASE COUNT 13 a 5 c 4 g 6 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 258; Length 28;
Best Local Similarity 85.7%; Pred. No. 1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 14
   |||||
Db 17 TATGCAGTTGCCAT 4

RESULT 10
LOCUS AA806226/c 40 bp mRNA EST 12-FEB-1998
DEFINITION oe29f12.sl NCIC_GAP_Pr25 Homo sapiens cDNA clone IMAGE:1409999 3'
similar to gb:M33552 LYMPHOCYTE-SPECIFIC PROTEIN LSPI (HUMAN);,
mRNA sequence.
ACCESSION AA806226
VERSION AA806226.1 GI:2874976
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```


REFERENCE 1 (bases 1 to 40)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Suzanne L. Topalian, M.D., Robert K. Bright, Ph.D.
 CDNA Library Preparation: Stratagene, Inc.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 1.

FEATURES

source

1..40

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1409999"
 /clone_lib="NCI-CGAP_Pr25"
 /tissue_type="epithelium (cell line)"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Organ: prostate; Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: oligo dt. Normal prostate epithelial cell line (HPV immortalized). 5' adaptor sequence: 5' GAATTCGCACGAC 3' 3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."
 7 a 6 c 11 g 16 t

BASE COUNT

7 a 6 c 11 g 16 t

Query Match 72.0%; Score 10.8; DB 12; Length 40;
 Best Local Similarity 85.7%; Pred. No. 1.1e+05;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 aagcagctgccatt 15

||||| |||||

Db 21 AAGCAGCCACCAT 8

RESULT 11

TA126G07Q 41 bp DNA GSS 13-DEC-2000

LOCUS T. brucei sheared genomic DNA clone 126g07, reverse sequence, genomic survey sequence.
 DEFINITION

ACCESSION AL463426

VERSION AL463426.1 GI:11833936

KEYWORDS GSS.

SOURCE Trypanosoma brucei.

ORGANISM Trypanosoma brucei

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma.

1 (bases 1 to 41)

Hall, N., Bowman, S., Lennard, N.J., Doggett, J., Atkin, R.,

Chillingworth, C., Ormond, D., Harris, B., El-Sayed, N., Hou, L.,

Melville, S.E., Rajadream, M.A. and Barrell, B.G.

Direct Submission

Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nhlesanger.ac.uk

COMMENT

Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is

described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

Email: neilsayed@tigr.org

Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES

source

1..41

/organism="Trypanosoma brucei"

/strain="TREU927"

/db_xref="taxon:5691"

/clone="126g07"

BASE COUNT 14 a 9 c 6 g 12 t

ORIGIN

Query Match 72.0%; Score 10.8; DB 258; Length 41;
 Best Local Similarity 85.7%; Pred. No. 1.1e+05;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 14

||||| |||||

Db 6 TAAGTAGATGCCAT 19

RESULT 12

AW249855/c

LOCUS

DEFINITION

mRNA sequence.

ACCESSION

AW249855

VERSION

AW249855.1 GI:6592848

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 42)

NIH-MGC <http://mgi.nci.nih.gov/>.

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1999)

Other ESTs: 2821534.5prime

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DTF CDNA Library Preparation: Ling

Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E.

Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing

project Clone distribution: MGC clone distribution information can

be found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html Base Calling / Quality

Scores: PHRED from University of Washington Genome Center. Vector

Trimming: cross_match from University of Washington Genome Center

PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley

Drosophila Genome Project. University of Washington Genome Center:

<http://www.genome.washington.edu> Low Quality Sequence: 10

contiguous PHRED high quality bases followed by vector sequence. Very

Low Quality Sequence: Trace file contained 42 contiguous distinct

peaks following vector sequence. Polyadenylation: Based upon the

presence of a XhoI site followed by a run of 14 or more T residues

at the beginning of the sequence, this cDNA insert was

polyadenylated.

Plate: LLCM7 row: A column: 23

High quality sequence stop: 10.

Location/Qualifiers

1..42

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2821534"

/clone_lib="NIH_MGC_7"

/tissue_type="small cell carcinoma"

/cell_line="MGC3"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 6 a 7 c 11 g 18 t

ORIGIN

Query Match 72.0%; Score 10.8; DB 113; Length 42;
Best Local Similarity 85.7%; Pred. No. 1.1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 aagcagctgccatt 15
||||||| |||
Db 26 AAGCAGCTCCCTT 13

RESULT 13
AA553766/c
LOCUS
DEFINITION
AA553766 44 bp mRNA EST 08-SEP-1997
similar to gb:M33552 LYMPHOCYTE-SPECIFIC PROTEIN LSP1 (HUMAN);
mRNA sequence.

ACCESSION
AA553766
VERSION
AA553766.1 GI:2324305
KEYWORDS
EST.
SOURCE
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 44)

REFERENCE
AUTHORS
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Elias Campo,
M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 865 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .44

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1014961"
/clone_lib="NCI-CCAP_Coll1"
/tissue_type="tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: colon; Vector: Bluescript SK-; Site_1: EcoRI
; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo
dT. Multiple colon tumors. 5' adaptor sequence: 5'
GAATTCGGCAGGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTATTTT 3' Average insert size: 1.1 kb."

BASE COUNT 9 a 7 c 11 g 17 t

ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 44;
Best Local Similarity 85.7%; Pred. No. 1.1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 aagcagctgccatt 15
||||||| |||
Db 21 AAGCAGCTCCCAAT 8

RESULT 14
AA600013/c
LOCUS
DEFINITION
AA600013 44 bp mRNA EST 12-MAY-1999
ag29h11.s1 Jia bone marrow stroma Homo sapiens cDNA clone
IMAGE:1091013 3' similar to gb:M33552 LYMPHOCYTE-SPECIFIC PROTEIN
LSP1 (HUMAN); mRNA sequence.

ACCESSION
AA600013
VERSION
AA600013.1 GI:2433638
KEYWORDS
EST.
SOURCE
human.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 44)

REFERENCE
AUTHORS
Jia, L., Robey, P., Young, M., Schuler, G., Powell, J., Yang, L., Lennon
, G., Hillier, L., Allen, M., Bowles, L., Geisel, S., Kucaba, T., Marra
, M., Martin, J., Steptoe, M., Tan, F., Theising, B., Bowers, Y., Wylie
, T., Waterston, R., Wilson, R. and Francomano, C.
WashU-MGB/NHGRI EST Project
Unpublished (1997)
Contact: Wilton RK / Jia L
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .44

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1091013"
/clone_lib="Jia bone marrow stroma"
/sex="mixed"
/tissue_type="bone marrow stroma"
/dev_stage="mixed"
/lab_host="XLI-Blue MRF"/SOLR"
/note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI;
mRNA made from human bone marrow stroma, cDNA made by
oligo-dT priming. Directionally cloned. Size-selected for
average insert size >0.5 kb. Library supplied by Dr. Libin
Jia (NHGRI)."

BASE COUNT 9 a 7 c 11 g 17 t

ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 44;
Best Local Similarity 85.7%; Pred. No. 1.1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 aagcagctgccatt 15
||||||| |||
Db 21 AAGCAGCCACCATT 8

RESULT 15
AA713885/c
LOCUS
DEFINITION
AA713885 44 bp mRNA EST 18-FEB-1998
nv79d06.s1 NCI-CCAP_Br4 Homo sapiens cDNA clone IMAGE:1236011
similar to gb:M33552 LYMPHOCYTE-SPECIFIC PROTEIN LSP1 (HUMAN);
mRNA sequence.

ACCESSION
AA713885
VERSION
AA713885.1 GI:2726159
KEYWORDS
EST.

BASE COUNT 9 a 7 c 11 g 17 t

ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 44;
Best Local Similarity 85.7%; Pred. No. 1.1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 aagcagctgccatt 15
||||||| |||
Db 21 AAGCAGCCACCATT 8

RESULT 15
AA713885/c
LOCUS
DEFINITION
AA713885 44 bp mRNA EST 18-FEB-1998
nv79d06.s1 NCI-CCAP_Br4 Homo sapiens cDNA clone IMAGE:1236011
similar to gb:M33552 LYMPHOCYTE-SPECIFIC PROTEIN LSP1 (HUMAN);
mRNA sequence.

ACCESSION
AA713885
VERSION
AA713885.1 GI:2726159
KEYWORDS
EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 44)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-femail.nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Kristina A. Cole, M.D.,
Ph.D. student, Rodrigo F. Chuqui, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 521 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 1.
FEATURES
source
1..44
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1236011"
/clone_lib="NCI_CGAP_Br4"
/sex="female"
/tissue_type="normal ductal tissue"
/lab_host="DH10B"
/note="Organ: breast; Vector: pAMP10; mRNA made from
normal breast ductal tissue, cDNA made by oligo-dt
priming. Non-directionally cloned. Size-selected on
agarose gel, average insert size 600 bp. Reference:
Krizman et al. (1996) Cancer Research 56:5380-5383."
BASE COUNT 9 a 6 c 11 g 18 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 11; Length 44;
Best Local Similarity 85.7%; Pred. No. 1.1e+05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 aagcagctgccatt 15
||||| |
Db 21 AAGCAGCCACCATT 8

Search completed: October 2, 2001, 15:01:12
Job time: 10841 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:53 ; Search time 3339.34 Seconds
(without alignments)
69.480 Million cell updates/sec

Title: US-09-757-100B-31

Perfect score: 15

Sequence: 1 taccagtgagtct 15

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 423528

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pl1:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_om:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vil:*
- 59: gb_v12:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rol:*
- 95: gb_ro2:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	12.4	82.7	21	10	I30544
C 2	12.4	82.7	22	9	AR067320
C 3	11.8	78.7	34	9	AR091705
C 4	11.8	78.7	48	9	AR075823
C 5	11.8	78.7	48	10	E30456
C 6	11.4	76.0	18	9	A34802
C 7	11.4	76.0	30	10	I36150
C 8	11	73.3	20	9	AR060552
					I30544 Sequence 7
					AR067320 Sequence
					AR091705 Sequence
					AR075823 Sequence
					E30456 Method for
					A34802 HSV probe.
					I36150 Sequence 34
					AR060552 Sequence

```

9      11 73.3 20 10 AX069241 Sequence
10     11 73.3 28 9  AR003365 Sequence
11     11 73.3 28 9  AR067869 Sequence
12     10.8 72.0 21 9  AR043988 Sequence
13     10.8 72.0 21 9  AR043989 Sequence
14     10.8 72.0 21 9  AR073521 Sequence
15     10.8 72.0 21 9  AR073522 Sequence
16     10.8 72.0 21 9  AR073522 Sequence
17     10.8 72.0 21 9  AX022131 Sequence
18     10.8 72.0 21 9  AX022132 Sequence
19     10.8 72.0 21 10 193392 Sequence 78
20     10.8 72.0 21 10 193393 Sequence 79
21     10.8 72.0 26 9  AR013888 Sequence
22     10.8 72.0 26 9  AR033842 Sequence
23     10.8 72.0 26 9  AR042502 Sequence
24     10.8 72.0 26 9  AR058382 Sequence
25     10.8 72.0 26 9  AR089208 Sequence
26     10.8 72.0 29 10 175320 Sequence 69
27     10.8 72.0 36 94 M0SIGHID2
28     10.8 72.0 37 9  AX010786 Sequence
29     10.8 72.0 40 9  AR013900 Sequence
30     10.8 72.0 40 9  AR033854 Sequence
31     10.8 72.0 40 9  AR042514 Sequence
32     10.8 72.0 40 9  AR058394 Sequence
33     10.8 72.0 40 9  AR088220 Sequence
34     10.8 72.0 42 9  A36516 Sequence 57
35     10.8 72.0 42 9  AR013872 Sequence
36     10.8 72.0 42 9  AR033826 Sequence
37     10.8 72.0 42 9  AR042486 Sequence
38     10.8 72.0 42 9  AR058366 Sequence
39     10.8 72.0 42 9  AR080149 Sequence
40     10.8 72.0 42 9  AR088192 Sequence
41     10.8 72.0 44 9  AR035220 Sequence
42     10.8 72.0 50 10 143338 Sequence 6
43     10.8 72.0 50 10 143338 Sequence 6
44     10.4 69.3 15 9  AX018749 Sequence
45     10.4 69.3 16 9  AX015626 Sequence

```

ALIGNMENTS

```

RESULT 1
I30544/c
LOCUS      I30544      21 bp      DNA
DEFINITION Sequence 7 from patent US 5580969.
ACCESSION  I30544
VERSION     I30544.1 GI:1821335
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 21)
AUTHORS     Hoke,G.D., Bradley,M.O., Williams,T.J. and Lee,C.
TITLE       Antisense oligonucleotides directed against human ICAM-1 RNA
JOURNAL     Patent: US 5580969-A 7 03-DEC-1996;
FEATURES    Location/Qualifiers
            1..21
            source
            /organism="unknown"

BASE COUNT      5 a      9 c      2 t
ORIGIN

Query Match      82.7%; Score 12.4; DB 10; Length 21;
Best Local Similarity 92.9%; Pred. No. 5.4e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 taccagggtgagtc 14
Db 18 TTCCAGGTGAGTC 5

RESULT 2

```

```

AR067320/c
LOCUS      AR067320      22 bp      DNA
DEFINITION Sequence 668 from patent US 5851760.
ACCESSION  AR067320
VERSION     AR067320.1 GI:5998542
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 22)
AUTHORS     Evans,G.A. and Smith,M.W.
TITLE       Method for generation of sequence sampled maps of complex genomes
JOURNAL     Patent: US 5851760-A 668 22-DEC-1998;
FEATURES    Location/Qualifiers
            1..22
            source
            /organism="unknown"

BASE COUNT      8 a      5 c      5 t
ORIGIN

Query Match      82.7%; Score 12.4; DB 9; Length 22;
Best Local Similarity 92.9%; Pred. No. 5.4e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 accagggtgagtc 15
Db 14 ATCCAGGTGAGTCT 1

RESULT 3
AR091705
LOCUS      AR091705      34 bp      DNA
DEFINITION Sequence 15 from patent US 5994505.
ACCESSION  AR091705
VERSION     AR091705.1 GI:10018459
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 34)
AUTHORS     Ting,J.Pan-Yung and Chin,K.
TITLE       Forms of class II MHC transactivator (CIITA)
JOURNAL     Patent: US 5994505-A 15 30-NOV-1999;
FEATURES    Location/Qualifiers
            1..34
            source
            /organism="unknown"

BASE COUNT      7 a     11 c     10 g      6 t
ORIGIN

Query Match      78.7%; Score 11.8; DB 9; Length 34;
Best Local Similarity 86.7%; Pred. No. 1.3e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taccagggtgagtc 15
Db 16 TACCAGGTGAGTCT 30

RESULT 4
AR075823/c
LOCUS      AR075823      48 bp      DNA
DEFINITION Sequence 6 from patent US 5958700.
ACCESSION  AR075823
VERSION     AR075823.1 GI:10002569
KEYWORDS   .
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 48)
AUTHORS     Nadeau,J.G., Pitner,J.Bruce, Linn,C.Preston and Schram,J.L.
TITLE       Detection of nucleic acids by fluorescence quenching
JOURNAL     Patent: US 5958700-A 6 28-SEP-1999;

```

```

FEATURES          Location/Qualifiers
source
1..48
BASE COUNT      15 a      8 c      10 g      15 t
ORIGIN

Query Match
Best Local Similarity 78.7%; Score 11.8; DB 9; Length 48;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taccaggtagtct 15
   ||| ||| ||| ||| |||
Db 41 TACTCAGATGAGTCT 27

RESULT 5
E30456/c
LOCUS      E30456      48 bp      DNA      PAT      07-FEB-2001
DEFINITION Method for detecting target nucleic acid sequence and
            oligonucleotide.
ACCESSION  E30456
VERSION     E30456.1 GI:13025613
KEYWORDS   JP 1999056380-A/6.
SOURCE     unidentified.
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 48)
AUTHORS   James,G.N.J.J., Pitona,C.P.R.R. and L.S.
TITLE     Method for detecting target nucleic acid sequence and
JOURNAL   Patent: JP 1999056380-A 6 02-MAR-1999;
          BECTON DICKINSON & CO
COMMENT   OS Unidentified
          PN JP 1999056380-A/6
          PD 02-MAR-1999
          PF 29-MAY-1998 JP 1998166141
          PR 30-MAY-1997 US 08/865,675
          PI JAMES G NADEAU,J BLUCE PITONA,C PRESTON RIN,JAMES L SHURAMU PC
          C12N15/09,C12Q1/68,G01N33/50,G01N33/566,C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
FH Key
FT source
          Location/Qualifiers
          1..48
          /organism="Unidentified".

FEATURES          Location/Qualifiers
source
1..48
BASE COUNT      15 a      8 c      10 g      15 t
ORIGIN

Query Match
Best Local Similarity 78.7%; Score 11.8; DB 10; Length 48;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taccaggtagtct 15
   ||| ||| ||| ||| |||
Db 41 TACTCAGATGAGTCT 27

RESULT 6
A34802/c
LOCUS      A34802      18 bp      DNA      PAT      16-JUL-1996
DEFINITION HSV probe.
ACCESSION  A34802
VERSION     A34802.1 GI:1568283
KEYWORDS   .
SOURCE     synthetic construct.
ORGANISM   synthetic construct.
REFERENCE  1 (bases 1 to 18)
AUTHORS   Renard,A. and Thiry,M.

TITLE
JOURNAL
PATENT: EP 0377349-A 20 11-JUL-1990;
EUROGENTEC S.A
FEATURES          Location/Qualifiers
source
1..18
BASE COUNT      5 a      4 c      7 g      2 t
ORIGIN

Recombinant polypeptides of the haemorrhagic septicemia virus in
fish
Patent: EP 0377349-A 20 11-JUL-1990;
EUROGENTEC S.A
FEATURES          Location/Qualifiers
source
1..18
BASE COUNT      5 a      4 c      7 g      2 t
ORIGIN

Query Match
Best Local Similarity 76.0%; Score 11.4; DB 9; Length 18;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 cccaggtagtct 15
   ||||| |||||
Db 14 CCCAGGTAGTCT 2

RESULT 7
I36150
LOCUS      I36150      30 bp      DNA      PAT      13-MAY-1997
DEFINITION Sequence 34 from patent US 5604131.
ACCESSION  I36150
VERSION     I36150.1 GI:2087374
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 30)
AUTHORS   Wadsworth,S., Snyder,B., Reddy,V.B. and Wei,C.
TITLE     cDNA-genomic DNA hybrid sequence encoding APP770 containing a
          genomic DNA insert of the KI and OX-2 regions
JOURNAL   Patent: US 5604131-A 34 18-FEB-1997;
          Location/Qualifiers
FEATURES          Location/Qualifiers
source
1..30
BASE COUNT      7 a      9 c      5 g      9 t
ORIGIN

Query Match
Best Local Similarity 92.3%; Score 11.4; DB 10; Length 30;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 cccaggtagtct 15
   ||||| |||||
Db 6 CCCAGGTAAGTCT 18

RESULT 8
AR060552
LOCUS      AR060552      20 bp      DNA      PAT      29-SEP-1999
DEFINITION Sequence 53 from patent US 5840693.
ACCESSION  AR060552
VERSION     AR060552.1 GI:5987002
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 20)
AUTHORS   Eriksson,U., Olofsson,B., Alitalo,K. and Pajusola,K.
TITLE     Vascular endothelial growth factor-B
JOURNAL   Patent: US 5840693-A 53 24-NOV-1998;
          Location/Qualifiers
FEATURES          Location/Qualifiers
source
1..20
BASE COUNT      5 a      4 c      6 g      5 t
ORIGIN

```

```
Query Match 73.3%; Score 11; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 cagggtgagctc 15
Db 8 CAGGTGAGTCT 18

RESULT 9
AX069241
LOCUS AX069241 20 bp DNA 25-JAN-2001
DEFINITION Sequence 52 from Patent WO0102594.
ACCESSION AX069241
VERSION AX069241.1 GI:12579122
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 20)
AUTHORS Graves,D.R., Thomsen,L., Catchpole,I.R. and Ford,M.J.
TITLE Dna constructs based on the eif4a gene promoter
JOURNAL Patent: WO 0102594-A 52 11-JAN-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
Location/Qualifiers
source 1..20
BASE COUNT 6 a 6 c 7 g 1 t
ORIGIN

Query Match 73.3%; Score 11; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 acccaggtgag 12
Db 5 ACCCAGGTGAG 15

RESULT 10
AR003365
LOCUS AR003365 28 bp DNA 04-DEC-1998
DEFINITION Sequence 4 from patent US 5744304.
ACCESSION AR003365
VERSION AR003365.1 GI:3964624
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
1 (bases 1 to 28)
AUTHORS Munford,R.S.
TITLE Inflammation-induced expression of a recombinant gene
JOURNAL Patent: US 5744304-A 4 28-APR-1998;
FEATURES
Location/Qualifiers
source 1..28
BASE COUNT 4 a 6 c 11 g 7 t
ORIGIN

Query Match 73.3%; Score 11; DB 9; Length 28;
Best Local Similarity 100.0%; Pred. No. 4.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 cccaggtgagt 13
Db 16 CCCAGGTGAGT 26

RESULT 11
```

```
Query Match 73.3%; Score 11; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 cagggtgagctc 15
Db 8 CAGGTGAGTCT 18

RESULT 9
AX069241
LOCUS AX069241 20 bp DNA 25-JAN-2001
DEFINITION Sequence 52 from Patent WO0102594.
ACCESSION AX069241
VERSION AX069241.1 GI:12579122
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 20)
AUTHORS Graves,D.R., Thomsen,L., Catchpole,I.R. and Ford,M.J.
TITLE Dna constructs based on the eif4a gene promoter
JOURNAL Patent: WO 0102594-A 52 11-JAN-2001;
GLAXO GROUP LIMITED (GB)
FEATURES
Location/Qualifiers
source 1..20
BASE COUNT 6 a 6 c 7 g 1 t
ORIGIN

Query Match 73.3%; Score 11; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 acccaggtgag 12
Db 5 ACCCAGGTGAG 15

RESULT 10
AR003365
LOCUS AR003365 28 bp DNA 04-DEC-1998
DEFINITION Sequence 4 from patent US 5744304.
ACCESSION AR003365
VERSION AR003365.1 GI:3964624
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
1 (bases 1 to 28)
AUTHORS Munford,R.S.
TITLE Inflammation-induced expression of a recombinant gene
JOURNAL Patent: US 5744304-A 4 28-APR-1998;
FEATURES
Location/Qualifiers
source 1..28
BASE COUNT 4 a 6 c 11 g 7 t
ORIGIN

Query Match 73.3%; Score 11; DB 9; Length 28;
Best Local Similarity 100.0%; Pred. No. 4.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 cccaggtgagt 13
Db 16 CCCAGGTGAGT 26

RESULT 11
```

```
AR067869
LOCUS AR067869 28 bp DNA
DEFINITION Sequence 4 from patent US 5851822.
ACCESSION AR067869
VERSION AR067869.1 GI:5999091
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
1 (bases 1 to 28)
AUTHORS Munford,R.S.
TITLE Inflammation-induced expression of a recombinant gene
JOURNAL Patent: US 5851822-A 4 22-DEC-1998;
FEATURES
Location/Qualifiers
source 1..28
BASE COUNT 4 a 6 c 11 g 7 t
ORIGIN

Query Match 73.3%; Score 11; DB 9; Length 28;
Best Local Similarity 100.0%; Pred. No. 4.5e+04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 cccaggtgagt 13
Db 16 CCCAGGTGAGT 26

RESULT 12
AR043988/c
LOCUS AR043988 21 bp DNA 29-SEP-1999
DEFINITION Sequence 78 from patent US 5817311.
ACCESSION AR043988
VERSION AR043988.1 GI:5965453
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
1 (bases 1 to 21)
AUTHORS Bazin,H. and Latinne,D.
TITLE Methods of inhibiting T-cell mediated immune responses with
LO-CD2a-specific antibodies
JOURNAL Patent: US 5817311-A 78 06-OCT-1998;
FEATURES
Location/Qualifiers
source 1..21
BASE COUNT 5 a 5 c 8 g 3 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 21;
Best Local Similarity 85.7%; Pred. No. 6.4e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taccaggtgagtc 14
Db 15 TCCTCAGGTGAGTC 2

RESULT 13
AR043989
LOCUS AR043989 21 bp DNA 29-SEP-1999
DEFINITION Sequence 79 from patent US 5817311.
ACCESSION AR043989
VERSION AR043989.1 GI:5965454
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
1 (bases 1 to 21)
AUTHORS Bazin,H. and Latinne,D.
TITLE Methods of inhibiting T-cell mediated immune responses with
```


JOURNAL LO-CD2a-specific antibodies
Patent: US 5817311-A 79 06-OCT-1998;
FEATURES Location/Qualifiers
source 1..21
BASE COUNT 3 a 8 c 5 g 5 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 21;
Best Local Similarity 85.7%; Pred. No. 6.4e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taccacagtgagtc 14
| | | | | | | | | | | | | | | | | | | | | |
Db 7 TCCTCAGGTGAGTC 20

RESULT 14
AR073521/c
LOCUS AR073521 21 bp DNA PAT 28-AUG-2000
DEFINITION Sequence 78 from patent US 5951983.
ACCESSION AR073521
VERSION AR073521.1 GI:10000285
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 21)
AUTHORS Bazin,H., Latinne,D., Kaplan,R., Kieber-Emmons,T., Postema,C.E. and White-Scharf,M.E.
TITLE Methods of inhibiting T cell mediated immune responses with humanized LO-CD2A-specific antibodies
JOURNAL Patent: US 5951983-A 78 14-SEP-1999;
FEATURES Location/Qualifiers
source 1..21
BASE COUNT 5 a 5 c 8 g 3 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 21;
Best Local Similarity 85.7%; Pred. No. 6.4e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taccacagtgagtc 14
| | | | | | | | | | | | | | | | | | | | | |
Db 15 TCCTCAGGTGAGTC 2

RESULT 15
AR073522
LOCUS AR073522 21 bp DNA PAT 28-AUG-2000
DEFINITION Sequence 79 from patent US 5951983.
ACCESSION AR073522
VERSION AR073522.1 GI:10000286
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 21)
AUTHORS Bazin,H., Latinne,D., Kaplan,R., Kieber-Emmons,T., Postema,C.E. and White-Scharf,M.E.
TITLE Methods of inhibiting T cell mediated immune responses with humanized LO-CD2A-specific antibodies
JOURNAL Patent: US 5951983-A 79 14-SEP-1999;
FEATURES Location/Qualifiers
source 1..21
BASE COUNT 3 a 8 c 5 g 5 t
ORIGIN

Query Match 72.0%; Score 10.8; DB 9; Length 21;
Best Local Similarity 85.7%; Pred. No. 6.4e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taccacagtgagtc 14
| | | | | | | | | | | | | | | | | | | | | |
Db 7 TCCTCAGGTGAGTC 20

Search completed: October 2, 2001, 15:56:54
Job time: 14178 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:12 ; Search time 10798.2 Seconds
(without alignments)
13.131 Million cell updates/sec

Title: US-09-757-100B-31

Perfect score: 15

Sequence: 1 taccagtgagtgct 15

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 76260

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 JOURNAL plasmid inserts
 COMMENT Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: gdunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0347 row: C column: 11
 Seq primer: CGTGTAAACGACGCGCAGT
 Class: plasmid ends
 High quality sequence stop: 25.
 Location/Qualifiers

FEATURES
 source
 1. .25
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0347C11"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 3 a 6 c 9 g 7 t
 ORIGIN
 Query Match 76.0%; Score 11.4; DB 245; Length 25;
 Best Local Similarity 92.3%; Pred. No. 3e+04;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 cccaggtgagctct 15
 Db 5 CTCAGGTGAGTCT 17

RESULT 2
 AZ303920/c

LOCUS AZ303920 32 bp DNA 29-SEP-2000
 DEFINITION LM0003F19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 Clone UUGC1M0003F19 R, DNA sequence.

ACCESSION AZ303920
 VERSION AZ303920.1 GI:10339372
 KEYWORDS GSS.
 SOURCE house mouse.

ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 32)
 DUNN,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	11.4	76.0	25	245	AZ506197	AZ506197 LM0347C11
2	11	73.3	32	241	AZ303920	AZ303920 LM0003F19
3	10.8	72.0	27	244	AZ495501	AZ495501 LM0331P21
4	10.8	72.0	43	249	AZ785625	AZ785625 LM0029L08
5	10.4	69.3	33	242	AZ382781	AZ382781 LM0140N08
6	10.4	69.3	34	103	A1873935	A1873935 wM44d08.x
7	10.4	69.3	35	243	AZ438946	AZ438946 LM0229L15
8	10.4	69.3	42	242	AZ377696	AZ377696 LM0132I11
9	10.4	69.3	46	190	W47560	W47560 zc35a01.r1
10	10.4	69.3	47	249	AZ795746	AZ795746 2M0051F07
11	10.4	69.3	48	250	AZ819513	AZ819513 2M0091L11
12	10.4	69.3	49	17	A1188352	A1188352 qd08R06.x
13	10.4	69.3	50	107	AU105506	AU105506 AU105506
14	10.4	69.3	50	107	AU105507	AU105507 AU105507
15	10.2	68.0	24	241	AZ303688	AZ303688 LM0003M11
16	10.2	68.0	43	244	AZ480568	AZ480568 LM0302M10
17	10.2	68.0	47	156	D38689	D38689 HUMC1264 Hu
18	10.2	68.0	48	258	BF740213	BF740213 hu57b11.x
19	10.2	68.0	49	168	AI024893	AI024893 ou52b04.x
20	10.2	68.0	50	15	A1024893	A1024893 ou52b04.x
21	10.2	68.0	50	107	AU106341	AU106341 AU106341
22	10.6	66.7	26	250	AZ827167	AZ827167 2M0103N22
23	10.6	66.7	34	250	AZ827876	AZ827876 2M0104M01
24	10.6	66.7	41	246	AZ580746	AZ580746 LM0369I11
25	10.6	66.7	48	188	T17575	T17575 mps v60 The
26	9.8	65.3	28	241	AZ313467	AZ313467 LM0029P06
27	9.8	65.3	28	242	AZ345677	AZ345677 LM0080B01
28	9.8	65.3	34	242	AZ349201	AZ349201 LM0086B08
29	9.8	65.3	37	102	A1862676	A1862676 wJ27f04.x
30	9.8	65.3	40	9	AA605733	AA605733 fai8d03.s
31	9.8	65.3	42	243	AZ427649	AZ427649 LM0209A15
32	9.8	65.3	43	244	AZ495842	AZ495842 LM0331L23
33	9.8	65.3	44	13	AA930614	AA930614 vY67a03.r
34	9.8	65.3	45	250	AZ28142	AZ28142 2M0104P24
35	9.8	65.3	46	23	A1663836	A1663836 uJ06a03.x
36	9.8	65.3	47	244	AZ476231	AZ476231 LM0294P14
37	9.8	65.3	49	241	AZ328192	AZ328192 LM0051A19
38	9.8	65.3	49	246	AZ607150	AZ607150 LM0429M24
39	9.8	65.3	50	107	AU102408	AU102408 AU102408
40	9.8	65.3	50	107	AU102658	AU102658 AU102658
41	9.8	65.3	50	107	AU106203	AU106203 AU106203
42	9.8	65.3	50	188	R88739	R88739 YP93c04.s1
43	9.4	62.7	23	241	AZ311201	AZ311201 LM0026F04
44	9.4	62.7	24	246	AZ588336	AZ588336 LM0396O06
45	9.4	62.7	25	241	AZ330576	AZ330576 LM0056E05

ALIGNMENTS

RESULT 1
 LOCUS AZ506197 25 bp DNA GSS 05-OCT-2000
 DEFINITION LM0347C11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 Clone UUGC1M0347C11 F, DNA sequence.
 AZ506197
 AZ506197.1 GI:10687513
 GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 25)
 DUNN,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.

and Wright,D.,Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0003 row: F column: 19
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 32.
 Location/Qualifiers
 1. 32

FEATURES

source
 1. 32
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_lib="UUC1M0003F19"
 /clone_lib="Mouse 10kb plasmid UUC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 5 a 10 c 9 g 8 t

BASE COUNT

ORIGIN

Query Match 73.3%; Score 11; DB 241; Length 32;
 Best Local Similarity 100.0%; Pred. No. 5.3e+04;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 acccaggtgag 12
 |||||
 Db 28 ACCCAGGTGAG 18

RESULT

3
 LOCUS AZ495501/c 27 bp DNA GSS 05-OCT-2000
 DEFINITION IM0331P21F Mouse 10kb plasmid UUC1M library Mus musculus genomic
 clone UUC1M0331P21 F, DNA sequence.
 ACCESSION AZ495501
 VERSION AZ495501.1 GI:10670983
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 27)
 REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D.,Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0331 row: P column: 21
 Seq primer: CGTGTAAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 27.
 Location/Qualifiers
 1. 27

FEATURES

source
 1. 27
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone_lib="UUC1M0331P21"
 /clone_lib="Mouse 10kb plasmid UUC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 11 a 3 c 8 g 5 t

BASE COUNT

ORIGIN

Query Match 72.0%; Score 10.8; DB 244; Length 27;
 Best Local Similarity 85.7%; Pred. No. 6.8e+04;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 taccacagtgagtc 14
 |||||
 Db 25 TACCCATGTGATTC 12

RESULT

4
 LOCUS AZ785625/c 43 bp DNA GSS 16-FEB-2001
 DEFINITION 2M0029L08R Mouse 10kb plasmid UUC1M library Mus musculus genomic
 clone UUC2M0029L08 R, DNA sequence.
 ACCESSION AZ785625
 VERSION AZ785625.1 GI:12922571
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 43)
 REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: gdunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0029 row: L column: 08
 Seq primer: CACACAGGAACACTATGACC
 Class: plasmid ends
 High quality sequence stop: 43.
 Location/Qualifiers

FEATURES

source
 1. .43
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC2M0029L08"
 /clone_lib="Mouse 10kb plasmid-UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain:XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 6 a 20 c 10 g 7 t

BASE COUNT
 ORIGIN

Query Match 72.0%; Score 10.8; DB 249; Length 43;
 Best Local Similarity 85.7%; Pred. No. 7e+04; Indels 0; Gaps 0;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 acccaggtgagct 15
 | | | | | | | | | |
 Db 15 AGCCAGGTGAGGCT 2

RESULT 5
 AZ382781/c
 LOCUS
 DEFINITION IM0140N08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0140N08 F, DNA sequence.
 ACCESSION AZ382781
 VERSION AZ382781.1 GI:10496481
 KEYWORDS GSS.
 SOURCE house musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 33)

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: gdunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0140 row: N column: 08
 Seq primer: CGTTGTAACGACGGCCAGT
 Class: plasmid ends
 High quality sequence stop: 33.
 Location/Qualifiers

FEATURES

source
 1. .33
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0140N08"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 4 a 10 c 8 g 11 t

BASE COUNT
 ORIGIN

Query Match 69.3%; Score 10.4; DB 242; Length 33;
 Best Local Similarity 91.7%; Pred. No. 1.2e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 taccacaggtgag 12
 | | | | | | | | | |
 Db 20 TACCACAGGTGCG 9

RESULT 6
 AI873935/c
 LOCUS
 DEFINITION Wm44d08.x1 NC1_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2438799 3' similar to TR:070396 070396 SIK SIMILAR PROTEIN. ;, mRNA sequence.
 ACCESSION AI873935
 VERSION AI873935.1 GI:5547984
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 34)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 Cloning Sequencing by: Washington University Genome Sequencing Center
 Cloning Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
 Seq primer: -40UP from Gibco
 High quality sequence stop: 1.

FEATURES

source

1. 34
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2438799"
 /clone_lib="NCI-CGAP_Ut4"
 /tissue_type="serous papillary carcinoma, high grade, 2
 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.48 kb. Life Technologies catalog #:
 11542-016"

BASE COUNT 7 a 8 c 7 g 12 t

ORIGIN

Query Match 69.3%; Score 10.4; DB 103; Length 34;
 Best Local Similarity 91.7%; Pred. NO. 1.2e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 acccaggtgagt 13

||| ||||| |||||

Db 34 ACACAGGTGAGT 23

RESULT 7

AZ438946 35 bp DNA GSS 03-OCT-2000
 LOCUS 1M022915F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 DEFINITION clone UUGCLM022915 F, DNA sequence.
 ACCESSION AZ438946
 VERSION AZ438946.1 GI:10562959
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 35)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mamoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

JOURNAL

COMMENT

Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: dunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
 Plate: 0229 row: L column: 15
 Seq primer: CGTTGTAACGACGCCAGT
 Class: plasmid ends
 High quality sequence stop: 35.

FEATURES

source

1. 35
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM022915"
 /clone_lib="Mouse 10kb plasmid UUGCLM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
 /note="Vector: pMD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 9 a 9 c 8 g 9 t

ORIGIN

Query Match 69.3%; Score 10.4; DB 243; Length 35;
 Best Local Similarity 91.7%; Pred. NO. 1.2e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 taccacaggtgag 12

||| ||||| |||||

Db 10 TACACAGGTGAG 21

RESULT 8

AZ377696 42 bp DNA GSS 02-OCT-2000
 LOCUS 1M0132111F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 DEFINITION clone UUGCLM013211 F, DNA sequence.
 ACCESSION AZ377696
 VERSION AZ377696.1 GI:10491396
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 42)
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
 Islam, H., Longacre, S., Mamoud, M., Meenen, E., Pedersen, T., Reilly,
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
 and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

JOURNAL

COMMENT

Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0132 row: 1 column: 11
 Seq primer: CGTTGTAAACGACGGCCACT
 Class: plasmid ends
 High quality sequence stop: 42.
 Location/Qualifiers

FEATURES

source
 1. .42
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0132111"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (g14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

12 a 12 c 9 g 9 t

Query Match 69.3%; Score 10.4; DB 242; Length 42;
 Best Local Similarity 91.7%; Pred. No. 1.2e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ccagggtgagtc 14
 |||||
 Db 2 CCCATGTGAGTC 13

RESULT 9

W47560/c
 LOCUS W47560 46 bp mRNA EST 11-OCT-1996
 DEFINITION zc35a01.r1 Soares senescent fibroblasts_NBHSF Homo sapiens CDNA
 clone IMAGE:324264 5' similar to PIR:S43417 S43417 RNA/DNA-binding
 Protein - mouse ;, mRNA sequence.

ACCESSION W47560.1 GI:1332239
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 46)
 AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
 The WashU-Merck EST Project
 Unpublished (1995)
 Contact: Wilson RK

TITLE

JOURNAL Washington University School of Medicine
 COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Insert length: 2909 Std Error: 0.00
 Seq primer: mob.REGA+ET
 High quality sequence stop: 1.
 Location/Qualifiers

FEATURES

source
 1. .46
 /organism="Homo sapiens"
 /db_xref="GDB:125577g"
 /db_xref="taxon:9606"
 /clone="IMAGE:324264"
 /clone_lib="Soares senescent fibroblasts_NBHSF"
 /tissue_type="senescent fibroblast"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Vector: pF7T3D (Pharmacia) with a modified polylinker V-TYPE: phagemid; Site_1: Not I; Site_2: Eco RI ; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTACCAATCTGAAGTGGAGCGCGGCATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pF7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 16 a 13 c 10 g 6 t 1 others

Query Match 69.3%; Score 10.4; DB 190; Length 46;
 Best Local Similarity 91.7%; Pred. No. 1.2e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 ccagggtgagtc 15
 |||||
 Db 36 CCGGGTGAGTCT 25

RESULT 10

AZ795746
 LOCUS AZ795746 47 bp DNA GSS 16-FEB-2001
 DEFINITION 2M0051F07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC2M0051F07 F, DNA sequence.

ACCESSION AZ795746
 VERSION AZ795746.1 GI:12943089
 KEYWORDS GSS.
 SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 47)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

TITLE

JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
 Plate: 0051 row: F column: 07
 Seq primer: CGTTGTAAACGACGGCCACT
 Class: plasmid ends
 High quality sequence stop: 47.

FEATURES
source

Location/Qualifiers
1. 47
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0051F07"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

6 a 17 c 14 g 10 t

Query Match 69.3%; Score 10.4; DB 249; Length 47;
Best Local Similarity 91.7%; Pred. No. 1.2e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ccaggtgagtc 14
||||| |||||
Db 24 CCCAAGTGAGTC 35

RESULT 11

A2819513 48 bp DNA 20-FEB-2001
LOCUS
DEFINITION
2M0091L11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0091L11 F, DNA sequence.
A2819513
VERSION
A2819513.1 GI:12989421
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus

REFERENCE
AUTHORS
1 (bases 1 to 48)
Dunn, B., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
TITLE
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL
COMMENT
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Genome Center
Rm 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0091 row: L column: 11
Seq primer: CGTGTAAACGACGGCCAGT
Class: plasmid ends

FEATURES

source

High quality sequence stop: 48.

Location/Qualifiers
1. 48
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0091L11"
/clone.lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

5 a 10 c 18 g 15 t

Query Match 69.3%; Score 10.4; DB 250; Length 48;
Best Local Similarity 91.7%; Pred. No. 1.2e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 ccaggtgagtc 15
||||| |||||
Db 35 CCAGCTGAGTCT 46

RESULT 12

A1188352 49 bp mRNA EST 28-OCT-1998
LOCUS
DEFINITION
qd08b06.xl Soares.placenta.8to9weeks.28bHP8to9w Homo sapiens CDNA
clone IMAGE:1723091 3' similar to TR:Q95341 Q95341 UBIQUINOL
CYTOCHROME C REDUCTASE ;, mRNA sequence.
A1188352
VERSION
A1188352.1 GI:3739561
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
1 (bases 1 to 49)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
COMMENT
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 1262 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers
1. 49
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1723091"

/clone_lib="Soares_placenta_8to9weeks_2NbHP8to9w"
 /dev_stage="two placentae: one from 8 weeks and another
 from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"

/note="Organ: placenta; Vector: pT7T3D.(Pharmacia) with a
 modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTACCAATCTCAAGTCGAGCGCGGATTTTGTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pT7T3 vector
 (Pharmacia). Library constructed by Bento Soares and
 M.Patima Bonaldo."

BASE COUNT 15 a 14 c 8 g 12 t
 ORIGIN

Query Match 69.3%; Score 10.4; DB 17; Length 49;
 Best Local Similarity 91.7%; Pred. No. 1.2e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 cccaggtgagtc 14
 ||||| |||||
 Db 3 CCCAGATGAGTC 14

RESULT 13

AU105506 50 bp mRNA EST 05-APR-2001
 LOCUS
 DEFINITION AU105506 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 HRC12806, mRNA sequence.

ACCESSION AU105506
 VERSION AU105506.1 GI:13555027
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
 ,K., Suyama,A. and Sugano,S.

TITLE Fine structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries
 Unpublished (2001)

JOURNAL Contact: Yutaka Suzuki
 COMMENT Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 ,S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

Location/Qualifiers
 1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HRC12806"

BASE COUNT 11 a 14 c 16 g 9 t
 ORIGIN

Query Match 69.3%; Score 10.4; DB 107; Length 50;
 Best Local Similarity 91.7%; Pred. No. 1.2e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 taccaggtgag 12
 | ||||| |||||
 Db 31 TCCCCAGGTGAG 42

RESULT 14

AU105507

LOCUS
 DEFINITION AU105507 50 bp mRNA EST 05-APR-2001
 AU105507 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 HRC12884, mRNA sequence.

ACCESSION AU105507
 VERSION AU105507
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
 ,K., Suyama,A. and Sugano,S.

TITLE Fine structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries
 Unpublished (2001)

JOURNAL Contact: Yutaka Suzuki
 COMMENT Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 ,S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES

Location/Qualifiers
 1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HRC12884"
 /clone_lib="Sugano Homo sapiens cDNA library"

BASE COUNT 11 a 14 c 17 g 8 t
 ORIGIN

Query Match 69.3%; Score 10.4; DB 107; Length 50;
 Best Local Similarity 91.7%; Pred. No. 1.2e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 taccaggtgag 12
 | ||||| |||||
 Db 31 TCCCCAGGTGAG 42

RESULT 15

AZ303688/c 24 bp DNA GSS 29-SEP-2000
 LOCUS
 DEFINITION 1M0003M11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0003M11 F, DNA sequence.

ACCESSION AZ303688
 VERSION AZ303688.1 GI:10339066
 KEYWORDS GSS.
 SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 24)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)

JOURNAL Contact: Robert B. Weiss
 COMMENT University of Utah Genome Center
 University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00
Plate: 0003 row: M column: 11
Seq primer: CGTTGTAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 24.

FEATURES

Location/Qualifiers

1..24
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0003M11"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 7 a 3 c 4 g 10 t
ORIGIN

Query Match 68.0%; Score 10.2; DB 241; Length 24;
Best Local Similarity 80.0%; Pred. No. 1.5e+05;
Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 taccacggtgagctc 15
|| |||| | ||||
Db 17 TATCCAGATAAGTCT 3

Search completed: October 2, 2001, 15:01:13
Job time: 10842 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:56:54 ; Search time 3339.34 Seconds
(without alignments)
69.480 Million cell updates/sec

Title: US-09-757-100B-33
Perfect score: 15
Sequence: 1 tggcttatcttcagt 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 423528
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pli:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_rod:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vil:*
59: gb_vl2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	12.4	82.7	27	9 AR042372	AR042372 Sequence
c 2	12.4	82.7	27	9 AR052280	AR052280 Sequence
c 3	12	80.0	21	10 AX095001	AX095001 Sequence
c 4	11.8	78.7	21	9 AX002270	AX002270 Sequence
c 5	11.8	78.7	21	10 E28251	E28251 Purificatio
6	11.8	78.7	30	9 AR012670	AR012670 Sequence
7	11.8	78.7	30	10 I25036	I25036 Sequence 10
8	11.8	78.7	30	10 I30498	I30498 Sequence 10

```

9      11.8  78.7  30  10  172560  Sequence 9
10     11.4  76.0  21  10  124623  Sequence 7
11     11.4  76.0  21  10  125220  Sequence 7
12     11.4  76.0  21  10  125220  Sequence 7
13     11.4  76.0  21  10  125220  Sequence 7
14     11.4  76.0  21  10  125220  Sequence 7
15     11.4  76.0  21  10  125220  Sequence 7
16     11.4  76.0  21  10  125220  Sequence 7
17     11.4  76.0  21  10  125220  Sequence 7
18     11.4  76.0  21  10  125220  Sequence 7
19     11.4  76.0  21  10  125220  Sequence 7
20     11.4  76.0  21  10  125220  Sequence 7
21     11.4  76.0  21  10  125220  Sequence 7
22     11.4  76.0  21  10  125220  Sequence 7
23     11.4  76.0  21  10  125220  Sequence 7
24     11.4  76.0  21  10  125220  Sequence 7
25     11.4  76.0  21  10  125220  Sequence 7
26     11.4  76.0  21  10  125220  Sequence 7
27     11.4  76.0  21  10  125220  Sequence 7
28     11.4  76.0  21  10  125220  Sequence 7
29     11.4  76.0  21  10  125220  Sequence 7
30     11.4  76.0  21  10  125220  Sequence 7
31     11.4  76.0  21  10  125220  Sequence 7
32     11.4  76.0  21  10  125220  Sequence 7
33     11.4  76.0  21  10  125220  Sequence 7
34     11.4  76.0  21  10  125220  Sequence 7
35     11.4  76.0  21  10  125220  Sequence 7
36     11.4  76.0  21  10  125220  Sequence 7
37     11.4  76.0  21  10  125220  Sequence 7
38     11.4  76.0  21  10  125220  Sequence 7
39     11.4  76.0  21  10  125220  Sequence 7
40     11.4  76.0  21  10  125220  Sequence 7
41     11.4  76.0  21  10  125220  Sequence 7
42     11.4  76.0  21  10  125220  Sequence 7
43     11.4  76.0  21  10  125220  Sequence 7
44     11.4  76.0  21  10  125220  Sequence 7
45     11.4  76.0  21  10  125220  Sequence 7

```

ALIGNMENTS

```

RESULT 1
LOCUS AR042372/c 27 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 11 from patent US 5811304.
ACCESSION AR042372
VERSION AR042372.1 GI:5962868
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Huang, S.
TITLE Nucleic acid molecules encoding retinoblastoma protein-interacting
JOURNAL zinc finger proteins
PATENT: US 5811304-A 11 22-SEP-1998;
FEATURES
Location/Qualifiers
source 1..27
BASE COUNT 9 a 7 c 9 g 2 t
ORIGIN

```

```

Query Match 82.7%; Score 12.4; DB 9; Length 27;
Best Local Similarity 92.9%; Pred. No. 1.5e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tggcttattcttcag 14
|||||
Db 27 TGGCTTTCTTCAG 14

```

```

RESULT 2
LOCUS AR052280/c 27 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 11 from patent US 5811008.
ACCESSION AR052280
VERSION AR052280.1 GI:5975644
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 27)
AUTHORS Huang, S.
TITLE Retinoblastoma protein-interacting zinc finger proteins
JOURNAL Patent: US 5811008-A 11 03-NOV-1998;
FEATURES
Location/Qualifiers
source 1..27
BASE COUNT 9 a 7 c 9 g 2 t
ORIGIN

Query Match 82.7%; Score 12.4; DB 9; Length 27;
Best Local Similarity 92.9%; Pred. No. 1.5e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tggcttattcttcag 14
|||||
Db 27 TGGCTTTCTTCAG 14

RESULT 3
LOCUS AX095001 21 bp DNA PAT 30-MAR-2001
DEFINITION Sequence 179 from Patent WO0118250.
ACCESSION AX095001
VERSION AX095001.1 GI:13511204
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 21)
AUTHORS Lander, E.S.; Gargill, M., Ireland, J.S., Bolck, S., Daley, G.G. and
McCarty, J.J.
TITLE Single nucleotide polymorphisms in genes
JOURNAL Patent: WO 0118250-A 179 15-MAR-2001;
WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source 1..21
BASE COUNT 3 a 6 c 4 g 7 t 1 others
ORIGIN

Query Match 80.0%; Score 12; DB 10; Length 21;
Best Local Similarity 85.7%; Pred. No. 2.6e+04;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tggcttattcttcag 14
|||||
Db 4 TGGCTGAYCTTCAG 17

RESULT 4
LOCUS AX002270/c 21 bp DNA PAT 10-MAR-2000
DEFINITION Sequence 7 from Patent EP0881288.
ACCESSION AX002270
VERSION AX002270.1 GI:7241961
KEYWORDS
SOURCE unidentified.

```

ORGANISM unidentified
REFERENCE 1 (bases 1 to 21)
AUTHORS Berglund,E.D. and Kirschbaum,B.D.
TITLE Purification of higher order transcription complexes from transgenic non-human animals
JOURNAL Patent: EP 0881288-A 7 02-DEC-1998;
HOECHST AG (DE)
FEATURES
source 1..21
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 6 a 4 c 8 g 3 t
ORIGIN
Query Match 78.7%; Score 11.8; DB 9; Length 21;
Best Local Similarity 86.7%; Pred. No. 3.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 tggcttatcttcagt 15
||||| |||||
Db 19 TGGCTTAACCTTCAGT 5
RESULT 5
E28251/c 21 bp DNA PAT 07-FEB-2001
LOCUS Purification of higher transcription complex from nonhuman
DEFINITION transgenic animal.
ACCESSION E28251
VERSION E28251.1 GI:13025285
KEYWORDS JP 1999004638-A/3.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 21)
AUTHORS Bernd,K.E.B. and Meisterernst,G.P.
TITLE Purification of higher transcription complex from nonhuman transgenic animal
JOURNAL Patent: JP 1999004638-A 3 12-JAN-1999;
HOECHST AG
COMMENT OS Unidentified
PN JP 1999004638-A/3
PD 12-JAN-1999
PF 26-MAY-1998 JP 1998144743
PR 26-MAY-1997 DE
PI BERND KIRSCHBAUM,ERIC BERGURUNTO,MICHAEL MEISTERERNST, PI GREG PORITTSU
PC A01K67/027,C12N15/09,C12P21/02//C12P21/08,(C12N15/09,C12R1:91), PC C12N15/00,
PC (C12N15/00,C12R1:91)
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT exon 1..21.
FEATURES
source 1..21
Location/Qualifiers
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 6 a 4 c 8 g 3 t
ORIGIN
Query Match 78.7%; Score 11.8; DB 10; Length 21;
Best Local Similarity 86.7%; Pred. No. 3.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 tggcttatcttcagt 15
||||| |||||
Db 19 TGGCTTAACCTTCAGT 5

RESULT 6
AR012670 30 bp DNA PAT 05-DEC-1998
LOCUS Sequence 9 from patent US 5763584.
DEFINITION AR012670
ACCESSION AR012670
VERSION AR012670.1 GI:3970988
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Godowski,P.J.
TITLE Receptor activation with hepatocyte growth factor agonists
JOURNAL Patent: US 5763584-A 9 09-JUN-1998;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
BASE COUNT 7 a 7 c 4 g 12 t
ORIGIN
Query Match 78.7%; Score 11.8; DB 9; Length 30;
Best Local Similarity 86.7%; Pred. No. 3.3e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 tggcttatcttcagt 15
||||| |||||
Db 7 TGACATATCTTCAGT 21
RESULT 7
I25036 30 bp DNA PAT 07-OCT-1996
LOCUS I25036
DEFINITION Sequence 10 from patent US 5547856.
ACCESSION I25036
VERSION I25036.1 GI:1604906
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Godowski,P.J., Lokker,N.A. and Mark,M.R.
TITLE Hepatocyte growth factor variants
JOURNAL Patent: US 5547856-A 10 20-AUG-1996;
FEATURES Location/Qualifiers
source 1..30
/organism="unknown"
BASE COUNT 7 a 7 c 4 g 12 t
ORIGIN
Query Match 78.7%; Score 11.8; DB 10; Length 30;
Best Local Similarity 86.7%; Pred. No. 3.3e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 tggcttatcttcagt 15
||||| |||||
Db 7 TGACATATCTTCAGT 21
RESULT 8
I30498 30 bp DNA PAT 06-FEB-1997
LOCUS I30498
DEFINITION Sequence 10 from patent US 5580963.
ACCESSION I30498
VERSION I30498.1 GI:1821289
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 30)
AUTHORS Godowski,P.J., Lokker,N.A. and Mark,M.R.

TITLE Single-chain hepatocyte growth factor variants
JOURNAL Patent: US 5580963-A 10 03-DEC-1996;
FEATURES Location/Qualifiers
Source 1..30
BASE COUNT 7 a 7 c 4 g 12 t
ORIGIN

Query Match 78.7%; Score 11.8; DB 10; Length 30;
Best Local Similarity 86.7%; Pred. No. 3.3e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tggcttatcttcagt 15
||| ||||| |||||
Db 7 TGACATATCTTCAGT 21

RESULT 9
LOCUS I72560 30 bp DNA PAT 03-APR-1998
DEFINITION Sequence 9 from patent US 5684136.
ACCESSION I72560
VERSION I72560.1 GI:3008699
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 30)
AUTHORS Godowski, P.J.
TITLE Chimeric hepatocyte growth factor (HGF) ligand variants
JOURNAL Patent: US 5684136-A 9 04-NOV-1997;
FEATURES Location/Qualifiers
Source 1..30
BASE COUNT 7 a 7 c 4 g 12 t
ORIGIN

Query Match 78.7%; Score 11.8; DB 10; Length 30;
Best Local Similarity 86.7%; Pred. No. 3.3e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tggcttatcttcagt 15
||| ||||| |||||
Db 7 TGACATATCTTCAGT 21

RESULT 10
LOCUS I24623 21 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 7 from patent US 5545527.
ACCESSION I24623
VERSION I24623.1 GI:1604493
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 21)
AUTHORS Stevens, J.K. and Dunn, J.M.
TITLE Method for testing for mutations in DNA from a patient sample
JOURNAL Patent: US 5545527-A 7 13-AUG-1996;
FEATURES Location/Qualifiers
Source 1..21
BASE COUNT 3 a 2 c 6 g 10 t
ORIGIN

Query Match 76.0%; Score 11.4; DB 10; Length 21;
Best Local Similarity 92.3%; Pred. No. 5.8e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 gcttatcttcagt 15
||||| |||||
Db 2 GCTTATGTTTCAGT 14

RESULT 11
LOCUS I25220 21 bp DNA PAT 07-OCT-1996
DEFINITION Sequence 7 from patent US 5550020.
ACCESSION I25220
VERSION I25220.1 GI:1605090
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 21)
AUTHORS Gallie, B.D., Dunn, J.M. and Stevens, J.K.
TITLE Method, reagents and kit for diagnosis and targeted screening for retinoblastoma
JOURNAL Patent: US 5550020-A 7 27-AUG-1996;
FEATURES Location/Qualifiers
Source 1..21
BASE COUNT 3 a 2 c 6 g 10 t
ORIGIN

Query Match 76.0%; Score 11.4; DB 10; Length 21;
Best Local Similarity 92.3%; Pred. No. 5.8e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 gcttatcttcagt 15
||||| |||||
Db 2 GCTTATGTTTCAGT 14

RESULT 12
LOCUS AX093520/c 21 bp DNA PAT 30-MAR-2001
DEFINITION Sequence 50 from Patent WO0118198.
ACCESSION AX093520
VERSION AX093520.1 GI:13509959
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 21)
AUTHORS Weissenbach, J. and Hazan, J.
TITLE Cloning, expression and characterisation of the spg4 gene responsible for the most frequent form of autosomal spastic paraplegia
JOURNAL Patent: WO 0118198-A 50 15-MAR-2001;
FEATURES Location/Qualifiers
Source 1..21
BASE COUNT 8 a 4 c 4 g 5 t
ORIGIN

Query Match 73.3%; Score 11; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ttatcttcagt 15
||||| |||||
Db 13 TTATCTTCAGT 3

RESULT 13
AR040915

LOCUS AR040915 24 bp DNA 29-SEP-1999
DEFINITION Sequence 14 from patent US 5811238.
ACCESSION AR040915
VERSION AR040915.1 GI:5961411
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 24)
AUTHORS Stemmer,W.P.C. and Cramer,A.
TITLE Methods for generating polynucleotides having desired
characteristics by iterative selection and recombination
JOURNAL Patent: US 5811238-A 14 22-SEP-1998;
FEATURES Location/Qualifiers
source 1. .24
BASE COUNT 5 a 6 c 1 g 12 t
ORIGIN
Query Match 73.3%; Score 11; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 cttatcttcag 14
|||||
Db 2 CTTATCTTCAG 12
RESULT 14
LOCUS AR051822 24 bp DNA 29-SEP-1999
DEFINITION Sequence 14 from patent US 5830721.
ACCESSION AR051822
VERSION AR051822.1 GI:5975186
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 24)
AUTHORS Stemmer,W.P.C. and Cramer,A.
TITLE DNA mutagenesis by random fragmentation and reassembly
JOURNAL Patent: US 5830721-A 14 03-NOV-1998;
FEATURES Location/Qualifiers
source 1. .24
BASE COUNT 5 a 6 c 1 g 12 t
ORIGIN
Query Match 73.3%; Score 11; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 cttatcttcag 14
|||||
Db 2 CTTATCTTCAG 12
RESULT 15
LOCUS I36212 24 bp DNA 13-MAY-1997
DEFINITION Sequence 14 from patent US 5605793.
ACCESSION I36212
VERSION I36212.1 GI:2086725
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 24)
AUTHORS Stemmer,W.P.C.
TITLE Methods for in vitro recombination
JOURNAL Patent: US 5605793-A 14 25-FEB-1997;

FEATURES
source Location/Qualifiers
1. .24
BASE COUNT 5 a 6 c 1 g 12 t
ORIGIN
Query Match 73.3%; Score 11; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 4 cttatcttcag 14
|||||
Db 2 CTTATCTTCAG 12
Search completed: October 2, 2001, 15:56:55
Job time: 14179 sec

1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *
10: gb_est10: *
11: gb_est11: *
12: gb_est12: *
13: gb_est13: *
14: gb_est14: *
15: gb_est15: *
16: gb_est16: *
17: gb_est17: *
18: gb_est18: *
19: gb_est19: *
20: gb_est20: *
21: gb_est21: *
22: gb_est22: *
23: gb_est23: *
24: gb_est24: *
25: gb_est25: *
26: gb_est26: *
27: gb_est27: *
28: gb_est28: *
29: gb_est29: *
30: gb_est30: *
31: gb_est31: *
32: gb_est32: *
33: gb_est33: *
34: gb_est34: *
35: gb_est35: *
36: gb_est36: *
37: gb_est37: *
38: gb_est38: *
39: gb_est39: *
40: gb_est40: *
41: gb_est41: *
42: gb_est42: *
43: gb_est43: *

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts (5000)
JOURNAL Published (in print)
COMMENT Contact: Robert B. Weiss
University of Utah
Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

FEATURES	Location/Qualifiers
source	1. .21

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C 1	11.8	78.7	21	249	AZ787920	2M0034M09
C 2	11.4	76.0	43	250	AZ830358	2M0109M04
C 3	10.8	72.0	23	240	AZ649569	1M0519N09
C 4	10.8	72.0	40	258	TA187A02Q	AL476274 T. brucei
C 5	10.4	69.3	28	16	AI118907	AI118907 uci5f07.r
C 6	10.4	69.3	37	13	AA906910	QJ70B12.s
C 7	10.4	69.3	37	14	AA946896	QJ53f01.s
C 8	10.4	69.3	38	156	D45802	HUMGS03019
C 9	10.4	69.3	42	106	AO10757	AU010757 AU010757
C 10	10.4	69.3	45	249	AZ772066	1M0574I08
C 11	10.4	69.3	46	241	AZ324419	1M0046520
C 12	10.4	69.3	46	249	AZ793395	2M0046D07
C 13	10.4	69.3	49	120	AW772664	xn02b05.y
C 14	10.4	69.3	50	107	AU103710	AU103710 AU103710
C 15	10.2	68.0	22	258	TA28BE04P	AI487174 T. brucei
C 16	10.2	68.0	37	23	A1664274	ue83g02.r
C 17	10.2	68.0	42	246	AZ593631	1M0405H14
C 18	10.2	68.0	44	13	AA883443	am25d01.s
C 19	10.2	68.0	47	172	BG030013	602297169
C 20	10.2	68.0	50	251	AZ919467	1006015E0
C 21	10	66.7	25	23	A1647975	uk38f04.x
C 22	10	66.7	36	250	AZ809373	2M0073A20
C 23	10	66.7	39	258	TA129B06P	AL463983 T. brucei
C 24	10	66.7	40	12	AA862574	oh44Q09.s
C 25	10	66.7	44	247	AZ660568	1M0538B15
C 26	10	66.7	45	18	A1309096	tb57g10.x
C 27	10	66.7	49	11	AA711423	yu25e09.r
C 28	9.8	65.3	27	189	T96372	ye49e01.sl
C 29	9.8	65.3	31	113	AW247904	2820039.3
C 30	9.8	65.3	31	250	AZ849227	2M0150M02
C 31	9.8	65.3	31	258	TA52H11Q	AL456059 T. brucei
C 32	9.8	65.3	32	243	AZ420867	1M0198F22
C 33	9.8	65.3	36	12	AA860352	AL456059 T. brucei
C 34	9.8	65.3	39	249	AZ767935	AL456059 T. brucei
C 35	9.8	65.3	42	249	AZ772691	1M0567E12
C 36	9.8	65.3	46	164	BE209004	SMOVMfCAR
C 37	9.8	65.3	49	258	TA264D06Q	AL486984 T. brucei
C 38	9.8	65.3	50	104	AJ237103	AJ237103 AJ237103
C 39	9.8	65.3	50	107	AU102434	AU102434 AU102434
C 40	9.8	65.3	50	107	AU102435	AU102435 AU102435
C 41	9.8	65.3	50	107	AU102436	AU102436 AU102436
C 42	9.8	65.3	50	107	AU102437	AU102437 AU102437
C 43	9.8	65.3	50	107	AU102439	AU102439 AU102439
C 44	9.8	65.3	50	107	AU102440	AU102440 AU102440
C 45	9.8	65.3	50	107	AU102441	AU102441 AU102441

RESULT	1			
AZ787920/c				
LOCUS		21 bp	DNA	GSS
DEFINITION		2M0034M09R Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC2M0034M09 R, DNA sequence.		
ACCESSION	AZ787920			
VERSION	AZ787920.1	GI:12927197		
KEYWORDS	GSS.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus			
REFERENCE	1 (bases 1 to 21)			
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A., and Wright,D. Weiss,R.			

TITLE
JOURNAL
COMMENT

and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000).
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0109 row: M column: 04
Seq primer: CACACAGGAAACAGTATGACC
Class: plasmid ends
High quality sequence stop: 43.

FEATURES

source

Location/Qualifiers

1. .43

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUCG2M0109M04"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114[gb|AF129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT

ORIGIN

7 a 10 c 8 g 18 t

Query Match 76.0%; Score 11.4; DB 250; Length 43;
Best Local Similarity 92.3%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tggcttatcttca 13
Db 8 TGGCTTATCTTTA 20

RESULT 3

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AZ649569/c
1M0519N09F Mouse 10kb plasmid UUGCLM library Mus musculus genomic clone UUGCLM0519N09 F, DNA sequence.
AZ649569
GSS.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 23)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Malmoud,M., Meenen,E., Pedersen,T., Reilly

TITLE
JOURNAL
COMMENT

M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000).
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0519 row: N column: 09
Seq primer: GGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 23.

FEATURES

Location/Qualifiers

1. .23

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGCLM0519N09"

/clone_lib="Mouse 10kb plasmid UUGCLM library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114[gb|AF129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 4 c 3 g 5 t

ORIGIN

Query Match 72.0%; Score 10.8; DB 247; Length 23;
Best Local Similarity 85.7%; Pred. No. 7.5e+04;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tggcttatcttca 14
Db 20 TGGCATATTTTCAG 7

RESULT 4

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TA187A02Q/c
TA187A02Q 40 bp DNA
T. brucei sheared genomic DNA clone 187a02, reverse sequence, genomic survey sequence.
AL476274
AL476274.1 GI:11841048
GSS.
Trypanosoma brucei.
Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
1 (bases 1 to 40)
Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,

TITLE
JOURNAL
COMMENT

Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L., Melville,S.E., Rajandream,M.A. and Barrell,B.G.
 Direct Submission
 Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and nhls@sanger.ac.uk
 Constructed at the Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from a cloned population of Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared to give a tight size distribution (4 kb). The v + i method used for the library construction is described in detail in Smith, H. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).
 Email: nelsayed@tigr.org
 Details of T. brucei sequencing at the Sanger Centre are available at http://www.sanger.ac.uk/projects/T_brucei/.

FEATURES
 source

1. .40
 Location/Qualifiers
 /organism="Trypanosoma brucei"
 /strain="TREU927"
 /db_xref="taxon:5691"
 /clone="187a02"

BASE COUNT 9 a 9 c 14 g 8 t

ORIGIN

Query Match 72.0%; Score 10.8; DB 258; Length 40;
 Best Local Similarity 85.7%; Pred. No. 8e+04;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttgcttcttcag 14

||||| |||||

Db 23 TGGGTCTCTTCAG 10

RESULT
 5
 AII18907/c

LOCUS

DEFINITION uc15f07.r1 Soares_mammary_gland_NBMWG Mus musculus cDNA clone IMAGE:1398085 5' similar to TR:Q13825 Q13825 AU-BINDING PROTEIN/ENOYL-COA HYDRATASE. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

28 bp mRNA EST 02-SEP-1998
 AII18907 Soares_mammary_gland_NBMWG Mus musculus cDNA clone uc15f07.r1
 IMAGE:1398085 5' similar to TR:Q13825 Q13825 AU-BINDING PROTEIN/ENOYL-COA HYDRATASE. ; mRNA sequence.
 AII18907
 AII18907.1 GI:3519231
 EST.
 house mouse.
 Mus musculus
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 28)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLM; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI-909801
 Trace considered overall poor quality
 Possible reversed clone: similarity on wrong strand
 Seq primer: -28m13 rev2 ET from Amersham
 High quality sequence stop: 1.

FEATURES
 source

Location/Qualifiers
 1. .28
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1398085"
 /clone_lib="Soares_mammary_gland_NBMWG"
 /sex="male"
 /tissue_type="mammary gland"
 /dev_stage="4 weeks"
 /lab_host="DH10B"

/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TCTTACCAATCTGAAGTGGAGCGCGCGAATGGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 9 a 4 c 10 g 5 t

ORIGIN

Query Match 69.3%; Score 10.4; DB 16; Length 28;
 Best Local Similarity 91.7%; Pred. No. 1.3e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 cttatcttcagt 15

||||| ||

Db 19 CTTATCTTCGT 8

RESULT
 6
 AA906910/c

LOCUS

DEFINITION AA906910 37 bp mRNA EST 09-JUN-1998
 o170b12.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1503647 3' similar to TR:Q33565 Q33565 EATRO 164 KINETOPLAST CR3. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

EMAIL

NOTE

TRACE

INSERT

SEQ

PRIMER

FROM

AMERSHAM

HIGH

QUALITY

SEQUENCE

STOP

1.

Location/Qualifiers

1. .37

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1503647"

/clone_lib="Soares_NFL_T_GBC_S1"

/lab_host="DH10B"

/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made in

26 a 2 c 4 g 5 t

```

Query Match      69.3%; Score 10.4; DB 13; Length 37;
Best Local Similarity 91.7%;
Pred. No. 1.3e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  tggcttatcttc 12
         || |||||
Db      34  TGTCTTATCTTC 23

```

RESULT	7	
AA946896		
LOCUS		
DEFINITION		
	AA946896	37 bp mRNA
	q053f01.s1	NCI-CCAP-KiD5 Homo sapiens cDNA clone IMAGE:159073 3'
		similar to TR:007708 007708 HYPOTHETICAL 23.7 KD PROTEIN. ; mRNA
		sequence.
ACCESSION	AA946896	
VERSION	AA946896.1	GI:3110291
KEYWORDS		
SOURCE		human.
		EST
		23-JUL-1998

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 37)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
 Insert length: 1106 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 1.

Query Match 69.38; Score 10.4; DB 14; Length 37;

```

Best Local Similarity 91.7%; Pred. NO. 1.3e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 cttatcttcagt 15
    |||||
Db 11 CTTATCTTCATT 22

```

RESULT	8
D45802/c	
LOCUS	D45802 38 bp mRNA EST 20-FEB-1995
DEFINITION	HUMGS03019 Human adult lung 3' directed MboI cDNA Homo sapiens CDNA 3', mRNA sequence.

VERSION	D45802.1	GI:662756
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1, (bases 1 to 38)	
TITLE	Itoh,K., Okubo,K., Yosii,J., Yokouchi,H. and Matsubara,K.	
JOURNAL	An expression profile of active genes in human lung	
MEDLINE	DNA Research 1, 279-287 (1994)	
COMMENT	95236275	
	Contact: Kohichi Itoh	
	Institute for Molecular and Cellular Biology	
	Osaka University	
	3-1, Yamadaoka, Suita, Osaka, 565, Japan	
	Tel: 06-877-5111 x3910	
	Fax: 06-877-1922.	

```

Query Match      69.3%; Score 10.4; DB 156; Length 38;
Best Local Similarity 91.7%; Pred. No. 1.3e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  gcttatcttcag 14
          | | | | | | | |
Db       20  GTTTATCTTCAG 9

```

RESULT	9	
LOCUS	AU010757	
DEFINITION	42 bp mRNA	EST
	AU010757 Schizosaccharomyces pombe late log phase cDNA	31-JUL-1998
	Schizosaccharomyces pombe cDNA clone spc10335, mRNA sequence.	
ACCESSION	AU010757	
VERSION	AU010757.1	GI:3347437
KEYWORDS	EST.	
SOURCE	fission yeast.	
ORGANISM	Schizosaccharomyces pombe	

Genome Research Group
National Institute of Radiological Sciences
9-1, Anagawa-4-chome, Inage-ku, Chiba 263-8555, Japan
Email: mori.myo@nirs.go.jp

FEATURES
source

Location/Qualifiers
1. .42
/organism="Schizosaccharomyces pombe"
/strain="972"
/db_xref="taxon:4896"
/clone_lib="spci0335"
/sex="h minus"
/note="Vector: M13mp19; The cDNA library of Schizosaccharomyces pombe was prepared by cloning cDNA into the SmaI site of M13mp19 DNA and the direction of DNA sequences was not always from 5' to 3'. The cDNA data of Schizosaccharomyces pombe are available for searching on the World Wide Web. (URL, <http://www.nirs.go.jp>)"
6 a 10 c 8 g 18 t

BASE COUNT
ORIGIN

Query Match 69.3%; Score 10.4; DB 106; Length 42;
Best Local Similarity 91.7%; Pred. No. 1.3e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 gcttatcttcag 14
||||| |||||
Db 2 GCTATCTTCAG 13

RESULT 10
AZ772066/c

LOCUS AZ772066 45 bp DNA GSS 16-FEB-2001
DEFINITION IM0574L08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0574L08 R, DNA sequence.

ACCESSION AZ772066
VERSION AZ772066.1 GI:12894990
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 45)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0574 row: L column: 08
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 45.

FEATURES
source

Location/Qualifiers
1. .45
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 11 a 7 c 13 g 14 t
ORIGIN

Query Match 69.3%; Score 10.4; DB 249; Length 45;
Best Local Similarity 91.7%; Pred. No. 1.4e+05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 gcttatcttcag 14
||||| |||||
Db 16 GCTTATGTTTCAG 5

RESULT 11
AZ324419

LOCUS AZ324419 46 bp DNA GSS 29-SEP-2000
DEFINITION IM0046B20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0046B20 F, DNA sequence.

ACCESSION AZ324419
VERSION AZ324419.1 GI:10380118
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 46)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0046 row: B column: 20
Seq primer: CGTGTAAACAGCGGCCAGT
Class: plasmid ends
High quality sequence stop: 46.

FEATURES
source

Location/Qualifiers
1. .46
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource

(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMDA2 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 14 a 14 c 6 g 12 t
ORIGIN

Query Match 69.3%; Score 10.4; DB 241; Length 46;
Best Local Similarity 91.7%; Pred. No. 1.4e+05; Length 46;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 4 cttatcttoagt 15
| | | | | | | | | |
Db 25 CATATCTTCAGT 36

RESULT 12

LOCUS A2793395 46 bp DNA 16-FEB-2001
DEFINITION 2M0046D07R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0046D07 R, DNA sequence.
ACCESSION A2793395
VERSION A2793395.1 GI:12938305
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Mus. 1 (bases 1 to 46)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0046 row: D column: 07
Seq primer: CACACAGGAACACCTATGACC
Class: plasmid ends
High quality sequence stop: 46.

FEATURES
source
1. 46
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0046D07"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMDA2nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMDA2 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 15 a 8 c 3 g 20 t
ORIGIN

Query Match 69.3%; Score 10.4; DB 249; Length 46;
Best Local Similarity 91.7%; Pred. No. 1.4e+05; Length 46;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 tggcttatcttc 12
| | | | | | | | | |
Db 19 TGGCTTACCTTC 30

RESULT 13

LOCUS AW772664 49 bp mRNA EST 04-MAY-2000
DEFINITION xn02b05.y1 NCI_CGAP_L15 Homo sapiens cDNA clone IMAGE:2692497 5', mRNA sequence.
ACCESSION AW772664
VERSION AW772664.1 GI:7704663
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 49)
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40RP from Gibco.

FEATURES
source
1. 49
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2692497"
/clone_lib="NCI_CGAP_L15"
/tissue_type="hepatic adenoma"
/lab_host="DH10B"
/note="Organ: liver; Vector: pCMV-SPORT4; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 0.8 kb."
BASE COUNT 22 a 14 c 6 g 7 t
ORIGIN

Query Match 69.3%; Score 10.4; DB 120; Length 49;
 Best Local Similarity 91.7%; Pred. No. 1.4e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ggcttatcttca 13
 | | | | | | | | | |
 Db 9 GCCTTACTTCA 20

RESULT 14
 AU103710/c
 LOCUS AU103710 50 bp mRNA EST 05-APR-2001
 DEFINITION AU103710 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
 HEP14727, mRNA sequence.

ACCESSION AU103710
 VERSION AU103710
 KEYWORDS AU103710.1 GI:13553231
 SOURCE EST.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki,Y., Tsunoda,T., Taira,H., Mizushima-Sugano,J., Sese,J., Hata
 ,H., Ota,T., Isogai,T., Tanaka,T., Nakamura,Y., Morishita,S., Okubo
 ,K., Suyama,A. and Sugano,S.

TITLE Fine structural analysis of transcription start sites of human
 mRNAs using full-length enriched and 5'-end enriched cDNA libraries

JOURNAL Unpublished (2001)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yszuki@ims.u-tokyo.ac.jp
 Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
 ,S. Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES
 source
 1..50
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEP14727"
 /clone_lib="Sugano Homo sapiens cDNA library"
 BASE COUNT 9 a 12 c 20 g 9 t
 ORIGIN

Query Match 69.3%; Score 10.4; DB 107; Length 50;
 Best Local Similarity 91.7%; Pred. No. 1.4e+05;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ggcttatcttca 13
 | | | | | | | | | |
 Db 31 GCCTTACTTCA 20

RESULT 15
 TA288E04P
 LOCUS TA288E04P 22 bp DNA GSS 13-DEC-2000
 DEFINITION T. brucei sheared genomic DNA clone 288e04, forward sequence,
 genomic survey sequence.
 ACCESSION AL487174
 VERSION AL487174.1 GI:11853828
 KEYWORDS GSS.

SOURCE Trypanosoma brucei.
 ORGANISM Trypanosoma brucei
 Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
 Trypanosoma.

REFERENCE 1 (bases 1 to 22)
 AUTHORS Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
 Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
 Melville,S.E., Rajandream,M.A. and Barrell,B.G.

TITLE Direct Submission
 JOURNAL Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing

project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
 Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
 nh@sanger.ac.uk
 Constructed at the Institute for Genomic Research (TIGR),
 Rockville, MD. Genomic DNA isolated from a cloned population of
 Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
 to give a tight size distribution (4 kb). The v + i method used for the library construction is
 described in detail in Smith, H. and Venter, J.C. (Making small
 insert libraries for whole genome shotgun sequencing projects. In
 Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Barrell, Oxford University Press, 1999).

Email: nelsayed@tigr.org
 Details of T. brucei sequencing at the Sanger Centre are available
 at http://www.sanger.ac.uk/Projects/T_brucei/.

FEATURES
 source
 1..22
 /organism="Trypanosoma brucei"
 /strain="TREU927"
 /db_xref="taxon:5691"
 /clone="288e04"
 BASE COUNT 4 a 6 c 4 g 8 t
 ORIGIN

Query Match 68.0%; Score 10.2; DB 258; Length 22;
 Best Local Similarity 80.0%; Pred. No. 1.6e+05;
 Matches 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 tggcttatcttcagt 15
 | | | | | | | | | |
 Db 7 TCGCTTATGTTCCGT 21

Search completed: October 2, 2001, 15:01:15
 Job time: 10844 sec

;
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-7017

Query Match 64.0%; Score 12.8; DB 6; Length 17;
Best Local Similarity 87.5%; Pred. No. 3e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 aagcagctgccattat 19
|||||
Db 17 AAGCAGCTGCCACCAT 2

RESULT 2

US-09-866-108-7018/c
; Sequence 7018, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 7018
; LENGTH: 17
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-7018

Query Match 64.0%; Score 12.8; DB 6; Length 17;
Best Local Similarity 87.5%; Pred. No. 3e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 aagcagctgccattat 19
|||||
Db 17 AAGCAGCTGCCACCAT 2

Qy 4 aagcagctgccattat 19
|||||
Db 16 AAGCAGCTGCCACCAT 1

RESULT 3

US-09-866-108-11909/c
; Sequence 11909, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11909
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11909

Query Match 64.0%; Score 12.8; DB 6; Length 25;
Best Local Similarity 87.5%; Pred. No. 3.2e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 aagcagctgccattat 19
|||||
Db 25 AAGCAGCTGCCACCAT 10

RESULT 4

US-09-866-108-11910/c
; Sequence 11910, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong

```
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2001-05-25
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2001-05-25
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeonica Sequence Listing Engine
; SEQ ID NO 11910
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11910

Query Match 64.0%; Score 12.8; DB 6; Length 25;
Best Local Similarity 87.5%; Pred. No. 3.2e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 aagcagctgccattat 19
Db 24 AAGCAGCTGCCACCAT 9

RESULT 5
US-09-866-108-11911/c
; Sequence 11911, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
```

```
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeonica Sequence Listing Engine
; SEQ ID NO 11911
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11911

Query Match 64.0%; Score 12.8; DB 6; Length 25;
Best Local Similarity 87.5%; Pred. No. 3.2e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 aagcagctgccattat 19
Db 23 AAGCAGCTGCCACCAT 8

RESULT 6
US-09-866-108-11912/c
; Sequence 11912, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 11912
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11912

Query Match 64.0%; Score 12.8; DB 6; Length 25;
Best Local Similarity 87.5%; Pred. No. 3.2e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 aagcagctgccattat 19
||||| ||||| ||
Db 22 AAGCAGCTGCCACCAT 7

RESULT 7
US-09-866-108-11913/c
; Sequence 11913, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: ACOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 11913
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-11913

Query Match 64.0%; Score 12.8; DB 6; Length 25;
Best Local Similarity 87.5%; Pred. No. 3.2e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 aagcagctgccattat 19
||||| ||||| ||
Db 21 AAGCAGCTGCCACCAT 6

RESULT 8
US-09-866-108-11914/c
; Sequence 11914, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: ACOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05

RESULT 11
US-09-866-

QY 4 aagcagctgccattat 20
||| ||| ||| ||| |||
Db 16 aaacagatgccagtatt 32

Db 49 GTAAGCGAAGCCATTA 33
||||| | |||||

Search completed: October 2, 2001, 16:55:02
Job time: 17665 sec

RESULT 14
US-60-278-561-6257/c
; Sequence 6257, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 6257
; LENGTH: 42
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: SNP00060057
; NAME/KEY: snp
; LOCATION: 26
; OTHER INFORMATION: 229652.27, 1901, A->G
US-60-278-561-6257

Query Match 61.0%; Score 12.2; DB 8; Length 42;
Best Local Similarity 82.4%; Pred. No. 7.5e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 taagcagctgccattat 19
| ||||| ||| |||
Db 30 TGAGCAGCTGGGATTAT 14

RESULT 15
US-09-681-508-1/c
; Sequence 1, Application US/09681508
; GENERAL INFORMATION:
; APPLICANT: Lin, Yun
; APPLICANT: Heil, James R
; APPLICANT: Jayasena, Sumedha D
; TITLE OF INVENTION: Aptamer Based Two-Site Binding Assay
; FILE REFERENCE: NEX 89
; CURRENT APPLICATION NUMBER: US/09/681,508
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,016
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 49
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Nucleic Acid Ligand
US-09-681-508-1

Query Match 61.0%; Score 12.2; DB 7; Length 49;
Best Local Similarity 82.4%; Pred. No. 7.7e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gtaagcagctgccatta 18

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 21:50:11 ; Search time 17695.9 Seconds
(without alignments)
16.701 Million cell updates/sec

Title: US-09-757-100B-9

Perfect score: 20

Sequence: 1 ttgtgtagtgtaggtat 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1415048 seqs, 7388405095 residues

Total number of hits satisfying chosen parameters: 4519004

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/US0612_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/US0613_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US098_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US099_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US099H_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US099I_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US099J_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US099K_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	100.0	20	1	PCT-US00-18999-9
2	20	100.0	20	17	US-09-377-310-9
3	20	100.0	20	29	US-09-757-100B-9
4	15	75.0	15	1	PCT-US00-18999-29
5	15	75.0	15	17	US-09-377-310-29
6	15	75.0	15	29	US-09-757-100B-29
7	15	75.0	20	1	PCT-US00-18999-8
8	15	75.0	20	17	US-09-377-310-8
9	15	75.0	20	29	US-09-757-100B-8
10	14.8	74.0	47	18	US-09-422-978-407
11	14.4	72.0	25	55	US-60-232-638-58459
12	14.4	72.0	25	55	US-60-232-638-16838
13	14.2	71.0	25	55	US-60-234-017-209206
14	13.8	69.0	20	18	US-09-486-816-26
15	13.8	69.0	25	55	US-60-232-638-50420
16	13.8	69.0	36	18	US-09-449-631-83
17	13.8	69.0	36	31	US-09-848-616-83
18	13.6	68.0	25	55	US-60-232-638-45369
19	13.6	68.0	50	14	US-09-040-266-1519
20	13.4	67.0	47	18	US-09-422-978-537
21	13.2	66.0	25	55	US-60-233-166-183276
22	13.2	66.0	25	55	US-60-233-166-365262
23	13.2	66.0	25	55	US-60-234-017-181973
24	13.2	66.0	25	55	US-60-234-017-181978
25	13.2	66.0	31	16	US-09-238-402-550
26	13	65.0	15	1	PCT-US00-18999-28
27	13	65.0	15	17	US-09-377-310-28
28	13	65.0	15	29	US-09-757-100B-28
29	13	65.0	25	55	US-60-232-638-44200
30	13	65.0	25	55	US-60-232-638-44223
31	12.8	64.0	25	55	US-60-232-638-121622
32	12.8	64.0	25	55	US-60-233-166-359242
33	12.8	64.0	25	55	US-60-233-166-365273
34	12.8	64.0	25	55	US-60-233-166-9949
35	12.8	64.0	25	55	US-60-233-620-84153
36	12.8	64.0	25	55	US-60-234-049-44209
37	12.6	63.0	25	26	US-09-660-220-36361
38	12.6	63.0	25	55	US-60-232-638-21996
39	12.6	63.0	25	55	US-60-232-638-45114
40	12.6	63.0	25	55	US-60-232-638-130842
41	12.6	63.0	25	55	US-60-232-638-27969
42	12.6	63.0	25	55	US-60-234-017-352215
43	12.6	63.0	25	55	US-60-234-049-85629
44	12.6	63.0	36	1	PCT-US99-18603-15
45	12.6	63.0	36	1	PCT-US99-18603-16

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

ALIGNMENTS

RESULT 1
 PCT-US00-18999-9
 ; Sequence 9, Application PC/TUS0018999
 ; GENERAL INFORMATION:
 ; APPLICANT: Isis Pharmaceuticals, Inc.
 ; APPLICANT: Monia, Brett P.
 ; APPLICANT: Gaarde, William A.
 ; APPLICANT: Nero, Pamela S.
 ; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
 ; FILE REFERENCE: ISPH-0476
 ; CURRENT FILING DATE: 2000-07-13
 ; CURRENT APPLICATION NUMBER: PCT/US00/18999
 ; PRIOR FILING DATE: 1999-08-19
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 20
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: antisense sequence
 PCT-US00-18999-9

Query Match 100.0%; Score 20; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttttgcctagctagggtat 20
 |||||
 DB 1 ttttgcctagctagggtat 20

RESULT 2
 US-09-377-310-9
 ; Sequence 9, Application US/09377310A
 ; GENERAL INFORMATION:
 ; APPLICANT: Monia, Brett P.
 ; APPLICANT: Gaarde, William A.
 ; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
 ; FILE REFERENCE: ISPH-0389
 ; CURRENT FILING DATE: 1999-08-19
 ; CURRENT APPLICATION NUMBER: US/09/377,310A
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 20
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: antisense sequence
 US-09-377-310-9

Query Match 100.0%; Score 20; DB 17; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttttgcctagctagggtat 20
 |||||
 DB 1 ttttgcctagctagggtat 20

RESULT 3
 US-09-757-100B-9
 ; Sequence 9, Application US/09757100B

; GENERAL INFORMATION:
 ; APPLICANT: Monia, Brett P.
 ; APPLICANT: Gaarde, William A.
 ; APPLICANT: Nero, Pamela S.
 ; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
 ; FILE REFERENCE: ISPH-0533
 ; CURRENT APPLICATION NUMBER: US/09/757,100B
 ; CURRENT FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 09/377,310
 ; PRIOR FILING DATE: 1999-08-19
 ; PRIOR APPLICATION NUMBER: PCT/US00/18999
 ; PRIOR FILING DATE: 2000-07-13
 ; NUMBER OF SEQ ID NOS: 44
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 20
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: antisense sequence
 US-09-757-100B-9

Query Match 100.0%; Score 20; DB 29; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttttgcctagctagggtat 20
 |||||
 DB 1 ttttgcctagctagggtat 20

RESULT 4
 PCT-US00-18999-29
 ; Sequence 29, Application PC/TUS0018999
 ; GENERAL INFORMATION:
 ; APPLICANT: Isis Pharmaceuticals, Inc.
 ; APPLICANT: Monia, Brett P.
 ; APPLICANT: Gaarde, William A.
 ; APPLICANT: Nero, Pamela S.
 ; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
 ; FILE REFERENCE: ISPH-0476
 ; CURRENT APPLICATION NUMBER: PCT/US00/18999
 ; CURRENT FILING DATE: 2000-07-13
 ; PRIOR APPLICATION NUMBER: 09/377,310
 ; PRIOR FILING DATE: 1999-08-19
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 29
 ; LENGTH: 15
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: antisense sequence
 PCT-US00-18999-29

Query Match 75.0%; Score 15; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ttgcctagctagg 17
 |||||
 DB 1 ttgcctagctagg 15

RESULT 5
 US-09-377-310-29
 ; Sequence 29, Application US/09377310A
 ; GENERAL INFORMATION:
 ; APPLICANT: Monia, Brett P.

; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-29

Query Match 75.0%; Score 15; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ttgctagatgctagg 17
|||||
Db 1 ttgctagatgctagg 15

RESULT 6
US-09-757-100B-29
; Sequence 29, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-29

Query Match 75.0%; Score 15; DB 29; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ttgctagatgctagg 17
|||||
Db 1 ttgctagatgctagg 15

RESULT 7
PCT-US00-18999-8
; Sequence 8, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476

; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-8

Query Match 75.0%; Score 15; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctatagtgctaggtat 20
|||||
Db 1 ctatagtgctaggtat 15

RESULT 8
US-09-377-310-8
; Sequence 8, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-8

Query Match 75.0%; Score 15; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctatagtgctaggtat 20
|||||
Db 1 ctatagtgctaggtat 15

RESULT 9
US-09-757-100B-8
; Sequence 8, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 8
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-8

Query Match      75.0%; Score 15; DB 29; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ctatagctaggtat 20
Db 1 ctatagctaggtat 15

RESULT 10
US-09-422-978-407
; Sequence 407, Application US/09422978
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
; FILE REFERENCE: GENSET.020CPI
; CURRENT APPLICATION NUMBER: US/09/422,978
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/298,850
; EARLIER FILING DATE: 1999-04-21
; EARLIER APPLICATION NUMBER: US 60/109,732
; EARLIER FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: US 60/082,614
; EARLIER FILING DATE: 1998-04-21
; NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 407
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-14678-75 : polymorphic base G or C
US-09-422-978-407

Query Match      74.0%; Score 14.8; DB 18; Length 47;
Best Local Similarity 88.9%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ttgctagatgctaggtat 20
Db 4 ttgcaagatgctaggaat 21

RESULT 11
US-60-232-638-58459/c
; Sequence 58459, Application US/60232638
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of Yeast
; FILE REFERENCE: 3110
; CURRENT APPLICATION NUMBER: US/60/232,638
; CURRENT FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 138410
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 58459
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; PUBLICATION INFORMATION:
```

```
; DATABASE ACCESSION NUMBER: SGD YER050C
US-60-232-638-58459

Query Match      72.0%; Score 14.4; DB 55; Length 25;
Best Local Similarity 93.8%; Pred. No. 3.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 tgctagatgctaggtat 19
Db 24 TGATAGATGCTAGGTA 9

RESULT 12
US-60-233-620-16838
; Sequence 16838, Application US/60233620
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of
; TITLE OF INVENTION: Arabidopsis thaliana
; FILE REFERENCE: 3116
; CURRENT APPLICATION NUMBER: US/60/233,620
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 131820
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16838
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AC005499
US-60-233-620-16838

Query Match      72.0%; Score 14.4; DB 55; Length 25;
Best Local Similarity 93.8%; Pred. No. 3.4e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 tgctagatgctaggtat 19
Db 1 tgctagaagctaggtat 16

RESULT 13
US-60-234-017-209206/c
; Sequence 209206, Application US/60234017
; GENERAL INFORMATION:
; APPLICANT: Mittmann, M
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of Mus
; TITLE OF INVENTION: musculus
; FILE REFERENCE: 3115
; CURRENT APPLICATION NUMBER: US/60/234,017
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 605887
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 209206
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AI504141
US-60-234-017-209206

Query Match      71.0%; Score 14.2; DB 55; Length 25;
Best Local Similarity 84.2%; Pred. No. 4.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ttgtctagatgctaggtat 20
Db 23 TTTGTTAGATGCTCGGTAT 5
```


Search completed: October 2, 2001, 21:50:12
Job time: 24525 sec

RESULT 14
US-09-486-816-26
; Sequence 26, Application US/09486816
; GENERAL INFORMATION:
; APPLICANT: HARADA, Hideyuki
; APPLICANT: FUJIWARA, Masanori
; APPLICANT: TAGASHIRA, Shuzo
; APPLICANT: OGAWA, Shinji
; APPLICANT: KATSUMATA, Takashi
; APPLICANT: NAKATSUKA, Masashi
; TITLE OF INVENTION: DNA Participating In Regulation of Expression of PEBP2alpha Ge
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0020-4685P
; CURRENT APPLICATION NUMBER: US/09/486,816
; CURRENT FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 254250/1997 Japan
; PRIOR FILING DATE: 1997-09-02
; PRIOR APPLICATION NUMBER: 299407/1997 Japan
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 114135/1998 Japan
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: PCT/JP98/03920
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 26
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Other nucleic acid, synthetic DNA
US-09-486-816-26

Query Match 69.0%; Score 13.8; DB 18; Length 20;
Best Local Similarity 88.2%; Pred. No. 6.6e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 ttgtctagatgctaggt 18
||| ||| ||| ||| ||| |||
Db 3 ttgtataaatgctaggt 19

RESULT 15
US-60-232-638-50420/C
; Sequence 50420, Application US/60232638
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of Yeast
; FILE REFERENCE: 3110
; CURRENT APPLICATION NUMBER: US/60/232,638
; CURRENT FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 138410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50420
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: SGD YDR240C
US-60-232-638-50420

Query Match 69.0%; Score 13.8; DB 55; Length 25;
Best Local Similarity 88.2%; Pred. No. 6.8e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 ttgtctagatgctaggt 18
||| ||| ||| ||| ||| |||
Db 22 TTTACTAGATGCTAGTT 6

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:55:01 ; Search time 2173.58 Seconds
(without alignments)
19.290 Million cell updates/sec

Title: US-09-757-100B-9

Perfect score: 20
Sequence: 1 ttttctagatgctaggtat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1895189 seqs, 1048201267 residues

Total number of hits satisfying chosen parameters: 299022

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*

- 1: /cgn2.6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2.6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2.6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2.6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2.6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2.6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 7: /cgn2.6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 8: /cgn2.6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	13.8	69.0	45	US-60-278-258-6107	Sequence 6107, Ap
C 2	13.2	66.0	31	US-09-574-376B-1127	Sequence 1127, Ap
C 3	12.8	64.0	31	US-09-801-274-1141	Sequence 1141, Ap
4	12.8	64.0	31	US-09-801-274-1801	Sequence 1801, Ap
5	12.8	64.0	31	US-09-817-879-5496	Sequence 5496, Ap
C 6	12.4	62.0	27	US-09-445-283B-47	Sequence 47, Appl
7	12.2	61.0	31	US-09-817-879-4811	Sequence 4811, Ap
8	12.2	61.0	31	US-09-817-879-5557	Sequence 5557, Ap
9	12.2	61.0	31	US-09-817-879-6920	Sequence 6920, Ap
10	12.2	61.0	31	US-09-817-879-9054	Sequence 9054, Ap
C 11	12	60.0	36	US-60-253-378-11076	Sequence 11076, A
12	11.8	59.0	31	US-09-708-690-16613	Sequence 16613, A
C 13	11.8	59.0	36	US-09-232-170-20	Sequence 20, Appl
14	11.8	59.0	38	US-09-840-424-773	Sequence 773, Appl
15	11.6	58.0	31	US-09-801-274-1751	Sequence 1751, Ap
C 16	11.6	58.0	49	US-60-253-654-12658	Sequence 12658, A
C 17	11.6	58.0	49	US-60-255-592-12658	Sequence 12658, A
18	11.4	57.0	15	US-09-895-007-96	Sequence 96, Appl
19	11.4	57.0	17	US-09-920-313-96	Sequence 313, Appl
C 20	11.4	57.0	17	US-09-817-879-2124	Sequence 2124, Ap
21	11.4	57.0	17	US-09-817-879-2431	Sequence 2431, Ap
C 22	11.4	57.0	20	US-09-827-383-1410	Sequence 1410, Ap
23	11.4	57.0	33	US-09-555-998B-15	Sequence 15, Appl
24	11.4	57.0	40	US-60-278-232-4276	Sequence 4276, Ap
C 25	11.4	57.0	40	US-60-253-457-8122	Sequence 8122, Ap

ALIGNMENTS

RESULT 1

US-60-278-258-6107/c
; Sequence 6107, Application US/60278258

; GENERAL INFORMATION:

; APPLICANT: Morris, MacDonald

; APPLICANT: Lal, Preeti

; APPLICANT: Diep, Dinh

; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using Polynucleotide Sequence Databases, and Single Nucleotide Polymorphisms Identified Thereby

; TITLE OF INVENTION: Polymorphisms Identified Thereby

; FILE REFERENCE: GX-0010-1 P

; CURRENT APPLICATION NUMBER: US/60/278,258

; CURRENT FILING DATE: 2001-03-23

; NUMBER OF SEQ ID NOS: 17730

; SOFTWARE: PERL Program

; SEQ ID NO 6107

; LENGTH: 45

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No: SNP00057112

; NAME/KEY: snp

; LOCATION: 20

; OTHER INFORMATION: 350403.1, 20, G->A

US-60-278-258-6107

Query Match 69.0%; Score 13.8; DB 8; Length 45;
Best Local Similarity 88.2%; Pred. No. 8.8e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 tgcctagatgctaggtat 20

|||||

Db 20 TGCTAGATGGTAGTAT 4

RESULT 2

US-09-574-376B-1127/c

; Sequence 1127, Application US/09574376B

; GENERAL INFORMATION:

; APPLICANT: Warrington, Janet

; APPLICANT: Shah, Nila

; APPLICANT: Gingers, Thomas Raymond

; APPLICANT: Affymetrix, Inc.

; TITLE OF INVENTION: Human Lymphoblast Polymorphisms

; FILE REFERENCE: 3229.2

; CURRENT APPLICATION NUMBER: US/09/574,376B

; CURRENT FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 1330
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1127
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLIC INFORMATION:
US-09-574-376B-1127

Query Match 66.0%; Score 13.2; DB 5; Length 31;
Best Local Similarity 75.0%; Pred. No. 1.7e+03;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttctgtagatgctaggtat 20
||| | : ||||| ||
Db 24 TTCTGCTCRATGCTAGAGAT 5

RESULT 3
US-09-801-274-1141
; Sequence 1141, Application US/09801274
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1141
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-1141

Query Match 64.0%; Score 12.8; DB 6; Length 31;
Best Local Similarity 77.8%; Pred. No. 2.8e+03;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ttgctagatgctaggtat 20
||| | : ||||| ||
Db 8 ttgccaaaygctcggtat 25

RESULT 4
US-09-801-274-1801
; Sequence 1801, Application US/09801274
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1801
; LENGTH: 31
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-801-274-1801

Query Match 64.0%; Score 12.8; DB 6; Length 31;
Best Local Similarity 77.8%; Pred. No. 2.8e+03;
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ttgctagatgctaggtat 20
||| | : ||||| ||
Db 8 ttgccaaaygctcggtat 25

RESULT 5
US-09-817-879-5496
; Sequence 5496, Application US/09817879
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Rel
; FILE REFERENCE: MH800-801-F
; CURRENT APPLICATION NUMBER: US/09/817,879
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9703
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5496
; LENGTH: 31
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: DNazyme
US-09-817-879-5496

Query Match 64.0%; Score 12.8; DB 6; Length 31;
Best Local Similarity 87.5%; Pred. No. 2.8e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 tgctagatgctaggtat 19
||| | : ||||| ||
Db 2 tgctagaggctagcta 17

RESULT 6
US-09-445-283B-47/C
; Sequence 47, Application US/09445283B
; GENERAL INFORMATION:
; APPLICANT: Maliga, Pal
; APPLICANT: Silhavy, Daniel
; APPLICANT: Sriraman, Priya
; TITLE OF INVENTION: Plastid Promoters for Transgene
; FILE REFERENCE: Rut 97-0097
; CURRENT APPLICATION NUMBER: US/09/445,283B
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: PCT/US98/11437
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/058,670
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/048,376
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 47
; LENGTH: 27
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-09-445-283B-47

Query Match 62.0%; Score 12.4; DB 6; Length 27;
Best Local Similarity 92.9%; Pred. No. 4.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ttgctagatgctag 16
| | | | | | | | | |
Db 17 TTGCTAGATTCTAG 4

RESULT 7

US-09-817-879-4811
; Sequence 4811, Application US/09817879
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to Hepatitis C Virus Infection
; FILE REFERENCE: MBH00-801-F
; CURRENT APPLICATION NUMBER: US/09/817,879
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9703
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4811
; LENGTH: 31
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: DNazyme
US-09-817-879-4811

Query Match 61.0%; Score 12.2; DB 6; Length 31;
Best Local Similarity 82.4%; Pred. No. 5.8e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ttgctagatgctagta 19
| | | | | | | | | |
Db 1 ttgctagagcctagta 17

RESULT 8

US-09-817-879-5557
; Sequence 5557, Application US/09817879
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to Hepatitis C Virus Infection
; FILE REFERENCE: MBH00-801-F
; CURRENT APPLICATION NUMBER: US/09/817,879
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9703
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5557
; LENGTH: 31
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: DNazyme
US-09-817-879-5557

Query Match 61.0%; Score 12.2; DB 6; Length 31;
Best Local Similarity 82.4%; Pred. No. 5.8e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ttgctagatgctagta 19
| | | | | | | | | |
Db 1 ttgctagagcctagta 17

RESULT 9

US-09-817-879-6920
; Sequence 6920, Application US/09817879
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to Hepatitis C Virus Infection
; FILE REFERENCE: MBH00-801-F
; CURRENT APPLICATION NUMBER: US/09/817,879
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9703
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6920
; LENGTH: 31
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: DNazyme
US-09-817-879-6920

Query Match 61.0%; Score 12.2; DB 6; Length 31;
Best Local Similarity 82.4%; Pred. No. 5.8e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ttgctagatgctagta 19
| | | | | | | | | |
Db 1 ttgctagagcctagta 17

RESULT 10

US-09-817-879-9054
; Sequence 9054, Application US/09817879
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to Hepatitis C Virus Infection
; FILE REFERENCE: MBH00-801-F
; CURRENT APPLICATION NUMBER: US/09/817,879
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9703
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9054
; LENGTH: 31
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: DNazyme
US-09-817-879-9054

Query Match 61.0%; Score 12.2; DB 6; Length 31;
Best Local Similarity 82.4%; Pred. No. 5.8e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ttgctagatgctagta 19
| | | | | | | | | |
Db 1 ttgctagagcctagta 17

RESULT 11

US-60-253-378-11076/c
; Sequence 11076, Application US/60253378
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J
; TITLE OF INVENTION: Polynucleotides isolated from plants and methods for their use.
; FILE REFERENCE: 1054P3

RESULT 13
US-09-232-170-20/c
Sequence 20, Application US/09232170
; GENERAL INFORMATION:
; APPLICANT: della-cioppa, Guy
; APPLICANT: Erwin, Robert L.
; APPLICANT: Fitzmaurice, Wayne P.
; APPLICANT: Hanley, Kathleen M.
; APPLICANT: Kumagai, Monto H.
; APPLICANT: Lendbo, John A.
; APPLICANT: McGeer, David R.

RESULT 15
US-09-801-274-1751
; Sequence 1751, Application US/09801274
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.

; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1751
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-1751

Query Match 58.0%; Score 11.6; DB 6; Length 31;
Best Local Similarity 70.0%; Pred. No. 1.2e+04;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
Qy 1 ttgtgctagatgctaggtat 20
||| ||| ||| |||
Db 8 ttctctasttgcaatgtat 27

Search completed: October 2, 2001, 16:55:01
Job time: 17664 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 21:50:10 ; Search time 17695.9 Seconds
(without alignments)
16.701 Million cell updates/sec

Title: US-09-757-100B-8

Perfect score: 20

Sequence: 1 ctatagctgtagtctgtc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1415048 seqs, 7388405095 residues

Total number of hits satisfying chosen parameters: 4519004

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/US0612_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/US0613_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/US0614_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US098_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US098E_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US098F_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US098G_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US098H_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US098I_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US098J_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US098K_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US098L_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	20	100.0	20	1	PCT-US00-18999-8	Sequence 8, Appli
2	20	100.0	20	17	US-09-377-310-8	Sequence 8, Appli
3	20	100.0	20	29	US-09-757-100B-8	Sequence 8, Appli
4	15.8	79.0	25	55	US-60-233-620-73486	Sequence 73486, A
5	15.8	79.0	25	55	US-60-233-620-73499	Sequence 73499, A
6	15.4	77.0	47	53	US-60-216-745-2764	Sequence 2764, Ap
7	15.2	76.0	25	55	US-60-233-620-101590	Sequence 101590, A
8	15	75.0	15	1	PCT-US00-18999-28	Sequence 28, Appl
9	15	75.0	15	17	US-09-377-310-28	Sequence 28, Appl
10	15	75.0	15	29	US-09-757-100B-28	Sequence 28, Appl
11	15	75.0	20	1	PCT-US00-18999-9	Sequence 9, Appli
12	15	75.0	20	17	US-09-377-310-9	Sequence 9, Appli
13	15	75.0	20	29	US-09-757-100B-9	Sequence 9, Appli
14	14.2	71.0	25	26	US-09-660-220-11883	Sequence 11883, A
15	14.2	71.0	25	26	US-09-660-220-48863	Sequence 48863, A
16	13.8	69.0	25	55	US-60-232-638-99167	Sequence 99167, A
17	13.6	68.0	21	18	US-09-422-978-8159	Sequence 8159, Ap
18	13.6	68.0	25	55	US-60-233-166-175158	Sequence 175158, A
19	13.6	68.0	25	55	US-60-233-166-218150	Sequence 218150, A
20	13.6	68.0	25	55	US-60-233-166-235863	Sequence 235863, A
21	13.6	68.0	25	55	US-60-233-166-321589	Sequence 321589, A
22	13.6	68.0	25	55	US-60-233-166-89443	Sequence 89443, A
23	13.4	67.0	25	55	US-60-234-017-94831	Sequence 94831, A
24	13.4	67.0	43	3	US-07-858-747A-17	Sequence 17, Appl
25	13.4	67.0	43	18	US-09-414-117-17	Sequence 17, Appl
26	13.2	66.0	25	26	US-09-660-220-48858	Sequence 48858, A
27	13.2	66.0	25	55	US-60-232-638-71961	Sequence 71961, A
28	13.2	66.0	25	55	US-60-232-638-123715	Sequence 123715, A
29	13.2	66.0	25	55	US-60-233-166-276832	Sequence 276832, A
30	13.2	66.0	25	55	US-60-233-166-359242	Sequence 359242, A
31	13.2	66.0	25	55	US-60-233-620-38163	Sequence 38163, A
32	13.2	66.0	25	55	US-60-233-620-65887	Sequence 65887, A
33	13.2	66.0	25	55	US-60-233-620-65897	Sequence 65897, A
34	13.2	66.0	25	55	US-60-234-017-4607	Sequence 4607, Ap
35	13.2	66.0	25	55	US-60-234-017-458673	Sequence 458673, A
36	13.2	66.0	25	55	US-60-234-017-458675	Sequence 458675, A
37	13.2	66.0	25	55	US-60-234-017-458675	Sequence 458675, A
38	13.2	66.0	25	55	US-60-234-017-458675	Sequence 458675, A
39	13	65.0	50	13	US-08-900-106-532	Sequence 532, App
40	12.8	64.0	25	55	US-60-232-638-58459	Sequence 58459, A
41	12.8	64.0	24	14	US-09-068-416-5	Sequence 5, Appli
42	12.8	64.0	25	26	US-09-660-220-14687	Sequence 14687, A
43	12.8	64.0	25	26	US-09-660-220-68406	Sequence 68406, A
44	12.8	64.0	25	55	US-60-233-620-13600	Sequence 13600, A
45	12.8	64.0	25	55	US-60-233-620-13607	Sequence 13607, A
						Sequence 13608, A

; CURRENT APPLICATION NUMBER: US/60/233,620
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 131820
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73499
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AC003672
US-60-233-620-73499

Query Match 79.0%; Score 15.8; DB 55; Length 25;
Best Local Similarity 89.5%; Pred. No. 3.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 tagatgctaggtagtctctc 20
||||| ||||||| |||
Db 6 tagaggctaggtagtcttc 24

RESULT 6
US-60-216-745-2764/c
; Sequence 2764, Application US/60216745
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Abderrahim, Hadi
; APPLICANT: Dufauré-Gare, Isabelle
; TITLE OF INVENTION: BIALLELIC MARKER MAPS FOR USE IN CONSTRUCTING A HIGH DENSITY...
; FILE REFERENCE: 84 US1 PRO
; CURRENT APPLICATION NUMBER: US/60/216,745
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 13665
; SOFTWARE: Patent.pm
; SEQ ID NO 2764
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 99-42676-348 : polymorphic base C or T
US-60-216-745-2764

Query Match 77.0%; Score 15.4; DB 53; Length 47;
Best Local Similarity 84.2%; Pred. No. 5.9e+02;
Matches 16; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 tagatgctaggtagtctc 20
||||| ||||||| |||
Db 40 TAGATGCTTAGTATCTTC 22

RESULT 7
US-60-233-620-101590/c
; Sequence 101590, Application US/60233620
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of
; FILE REFERENCE: 3116
; CURRENT APPLICATION NUMBER: US/60/233,620
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 131820
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 101590
; LENGTH: 25
; TYPE: DNA

; ORGANISM: Arabidopsis thaliana
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AF096373
US-60-233-620-101590

Query Match 76.0%; Score 15.2; DB 55; Length 25;
Best Local Similarity 85.0%; Pred. No. 6.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ctatgctaggtagtctctc 20
||||| ||||||| |||
Db 25 CTAGATGATAGGTACTCTC 6

RESULT 8
PCT-US00-18999-28
; Sequence 28, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-28

Query Match 75.0%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 agatgctaggtagtct 17
||||| ||||||| |||
Db 1 agatgctaggtagtct 15

RESULT 9
US-09-377-310-28
; Sequence 28, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-28

Query Match 75.0%; Score 15; DB 17; Length 15;

Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 agatgctaggtagtct 17
|||||
Db 1 agatgctaggtagtct 15

RESULT 10
US-09-757-100B-28
; Sequence 28, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR FILING DATE: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-28

Query Match 75.0%; Score 15; DB 29; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 agatgctaggtagtct 17
|||||
Db 1 agatgctaggtagtct 15

RESULT 11
PCT-US00-18999-9
; Sequence 9, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR FILING DATE: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-9

Query Match 75.0%; Score 15; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctatagtgctaggtagt 15
|||||
Db 6 ctatagtgctaggtagt 20

RESULT 12
US-09-377-310-9
; Sequence 9, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-9

Query Match 75.0%; Score 15; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctatagtgctaggtagt 15
|||||
Db 6 ctatagtgctaggtagt 20

RESULT 13
US-09-757-100B-9
; Sequence 9, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR FILING DATE: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-9

Query Match 75.0%; Score 15; DB 29; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctatagtgctaggtagt 15
|||||
Db 6 ctatagtgctaggtagt 20

RESULT 14

US-09-660-220-11883
; Sequence 11883, Application US/09660220
; GENERAL INFORMATION:
; APPLICANT: Mittmann et al.
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of Human
; FILE REFERENCE: 3102.1
; CURRENT APPLICATION NUMBER: US/09/660,220
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/164,973
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 140981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11883
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank D63881
US-09-660-220-11883

Query Match 71.0%; Score 14.2; DB 26; Length 25;
Best Local Similarity 84.2%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ctatgctctagggtatctgt 19
| | | | | | | | | | | | | | | | | | | | | |
Db 2 ctatgctctagggtatctgt 20

RESULT 15

US-09-660-220-48863/c
; Sequence 48863, Application US/09660220
; GENERAL INFORMATION:
; APPLICANT: Mittmann et al.
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of Human
; FILE REFERENCE: 3102.1
; CURRENT APPLICATION NUMBER: US/09/660,220
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/164,973
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 140981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48863
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo Sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank M14338
US-09-660-220-48863

Query Match 71.0%; Score 14.2; DB 26; Length 25;
Best Local Similarity 84.2%; Pred. No. 2.4e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 tagatgctagggtatctgtc 20
| | | | | | | | | | | | | | | | | | | | | |
Db 24 TAGATGGTAAGTATCGGTC 6

Search completed: October 2, 2001, 21:50:11
Job time: 24524 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:55:00 ; Search time 2173.58 Seconds
(without alignments)
19.290 Million cell updates/sec

Title: US-09-757-100B-8

Perfect score: 20

Sequence: 1 ctatagctaggtatctgtc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1895189 seqs, 1048201267 residues

Total number of hits satisfying chosen parameters: 299022

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq1:*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	12.8	64.0	38	7	US-09-840-424-773
2	12.2	61.0	45	8	US-60-278-258-6107
3	12	60.0	17	6	US-09-546-745A-3190
4	12	60.0	17	6	US-09-546-745A-3191
5	12	60.0	36	7	US-09-048-473-3
6	12	60.0	36	7	US-09-048-473-7
7	12	60.0	39	8	US-60-253-457-43320
8	11.8	59.0	17	6	US-09-866-108-2604
9	11.8	59.0	17	6	US-09-866-108-2605
10	11.8	59.0	17	6	US-09-866-108-2606
11	11.8	59.0	25	6	US-09-866-108-5533
12	11.8	59.0	25	6	US-09-866-108-5534
13	11.8	59.0	25	6	US-09-866-108-5535
14	11.8	59.0	25	6	US-09-866-108-5536
15	11.8	59.0	25	6	US-09-866-108-5537
16	11.8	59.0	25	6	US-09-866-108-5538
17	11.8	59.0	25	6	US-09-866-108-5539
18	11.8	59.0	25	6	US-09-866-108-5540
19	11.8	59.0	25	6	US-09-866-108-5541
20	11.8	59.0	25	6	US-09-866-108-5542
21	11.8	59.0	25	6	US-09-866-108-5543
22	11.8	59.0	50	4	US-08-798-074B-7431
23	11.8	59.0	50	4	US-08-798-074B-7431
24	11.8	59.0	50	4	US-08-798-074C-7431
25	11.8	59.0	50	7	US-09-839-976-3279

ALIGNMENTS

RESULT 1

US-09-840-424-773

; Sequence 773, Application US/09840424

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; APPLICANT: McCarthy, Sean A.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES DERIVED FROM A

; TITLE OF INVENTION: HUMAN MICROVASCULAR ENDOTHELIAL LIBRARY

; FILE REFERENCE: 1600.1010-002

; CURRENT APPLICATION NUMBER: US/09/840.424

; CURRENT FILING DATE: 2001-04-23

; PRIOR APPLICATION NUMBER: US 60/084,098

; PRIOR FILING DATE: 1998-05-04

; PRIOR APPLICATION NUMBER: US 60/123,523

; PRIOR FILING DATE: 1999-03-08

; PRIOR APPLICATION NUMBER: US 60/126,974

; PRIOR FILING DATE: 1999-03-30

; PRIOR APPLICATION NUMBER: US 09/304,649

; PRIOR FILING DATE: 1999-05-04

; NUMBER OF SEQ ID NOS: 1311

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 773

; LENGTH: 38

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(38)

; OTHER INFORMATION: n = A,T,C or G

US-09-840-424-773

Query Match 64.0%; Score 12.8; DB 7; Length 38;

Best Local Similarity 87.5%; Pred. No. 1.5e+03;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ctatagctaggtatc 16

||||||| |||

Db 16 ctatagtttagtgc 31

RESULT 2

US-60-278-258-6107/c

; Sequence 6107, Application US/60278258

; GENERAL INFORMATION:

; APPLICANT: Morris, MacDonald

; APPLICANT: Lal, Preeti

; APPLICANT: Diep, Dinh

Sequence 1033, Ap
Sequence 8033, Ap
Sequence 3189, Ap
Sequence 3, Appl
Sequence 32, Appl
Sequence 10, Appl
Sequence 26932, A
Sequence 87, Appl
Sequence 191, App
Sequence 22107, A
Sequence 16933, A
Sequence 17980, A
Sequence 18029, A
Sequence 2187, Ap
Sequence 34, Appl
Sequence 235, App
Sequence 3466, Ap
Sequence 2603, Ap
Sequence 2607, Ap
Sequence 100, App

; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0010-1.P
; CURRENT APPLICATION NUMBER: US/60/278,258
; NUMBER OF SEQ ID NOS: 2001-03-23
; SOFTWARE: PERL Program
; SEQ ID NO 6107
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: SNP00057112
; NAME/KEY: snp
; LOCATION: 20
; OTHER INFORMATION: 350403.1, 20, G->A
US-60-278-258-6107

Query Match 61.0%; Score 12.2; DB 8; Length 45;
Best Local Similarity 82.4%; Pred. No. 3.5e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ctagatgctaggtatct 17
||||| ||| ||| |
Db 18 CTAGATGGTAGTATCT 2

RESULT 3
US-09-546-745A-3190
; Sequence 3190, Application US/09546745A
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: Zwick, Michael
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Regulation of Repressor Genes using Nucleic Acid Molecules
; FILE REFERENCE: 237/193
; CURRENT APPLICATION NUMBER: US/09/546,745A
; CURRENT FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 7043
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3190
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Nucleic Acid
US-09-546-745A-3190

Query Match 60.0%; Score 12; DB 6; Length 17;
Best Local Similarity 58.3%; Pred. No. 3.8e+03;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 8 ctagggtatctgt 19
|:||||:|:|:
Db 3 cuagguaucugu 14

RESULT 4
US-09-546-745A-3191
; Sequence 3191, Application US/09546745A
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Blatt, Larry
; APPLICANT: Zwick, Michael
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; TITLE OF INVENTION: Regulation of Repressor Genes using Nucleic Acid Molecules

; FILE REFERENCE: 237/193
; CURRENT APPLICATION NUMBER: US/09/546,745A
; CURRENT FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 7043
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3191
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Nucleic Acid
US-09-546-745A-3191

Query Match 60.0%; Score 12; DB 6; Length 17;
Best Local Similarity 58.3%; Pred. No. 3.8e+03;
Matches 7; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 8 ctagggtatctgt 19
|:||||:|:|:
Db 1 cuagguaucugu 12

RESULT 5
US-09-048-473-3/c
; Sequence 3, Application US/09048473
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; Hayenga, Kirk J.
; Lawlis, Virgil B.
; Ward, Michael
; TITLE OF INVENTION: ASPARTIC PROTEINASE DEFICIENT
; FILAMENTOUS FUNGI
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International, Inc.
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304-1013
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/048,473
; FILING DATE: 26-Mar-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/214,237
; FILING DATE: 01-JUL-1992
; APPLICATION NUMBER: 07/931,123
; FILING DATE: 17-AUG-1992
; APPLICATION NUMBER: 08/345,018
; FILING DATE: 23-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J.
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: GC45-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-846-7620
; TELEFAX: 650-845-6504
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-048-473-3

Query Match 60.0%; Score 12; DB 7; Length 36;
Best Local Similarity 75.0%; Pred. No. 4.3e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ctatagctaggtatctgtc 20
||||| | ||| ||| ||| |||
Db 22 CTAGACTCGAGGATCCGTC 3

RESULT 6

US-09-048-473-7/c

; Sequence 7, Application US/09048473

; GENERAL INFORMATION:

; APPLICANT: Berkka, Randy M.

; Hayenga, Kirk J.

; Lawlis, Virgil B.

; Ward, Michael

; TITLE OF INVENTION: ASPARTIC PROTEINASE DEFICIENT

; FILAMENTOUS FUNGI

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genencor International, Inc.

; STREET: 925 Page Mill Road

; CITY: Palo Alto

; STATE: CA

; COUNTRY: US

; ZIP: 94304-1013

; COMPUTER READABLE FORM:

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/048,473

; FILING DATE: 26-Mar-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/214,237

; FILING DATE: 01-JUL-1992

; APPLICATION NUMBER: 07/931,123

; FILING DATE: 17-AUG-1992

; APPLICATION NUMBER: 08/345,018

; FILING DATE: 23-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Glaister, Debra J.

; REGISTRATION NUMBER: 33,888

; REFERENCE/DOCKET NUMBER: GC45-4

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-846-7620

; TELEFAX: 650-845-6504

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 36 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-048-473-7

Query Match 60.0%; Score 12; DB 7; Length 36;
Best Local Similarity 75.0%; Pred. No. 4.3e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ctatagctaggtatctgtc 20
||||| | ||| ||| ||| |||
Db 22 CTAGACTCGAGGATCCGTC 3

RESULT 7

US-09-253-457-43320/c

; Sequence 43320, Application US/60253457

; GENERAL INFORMATION:

; APPLICANT: Havukkala, Ilkka J

; TITLE OF INVENTION: Polynucleotides, isolated from plants

; FILE OF INVENTION: and methods for their use.

; FILE REFERENCE: 1054P2

; CURRENT APPLICATION NUMBER: US/60/253,457

; CURRENT FILING DATE: 2000-11-27

; NUMBER OF SEQ ID NOS: 48893

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 43320

; LENGTH: 39

; TYPE: DNA

; ORGANISM: Pinus radiata

US-60-253-457-43320

Query Match 60.0%; Score 12; DB 8; Length 39;

Best Local Similarity 75.0%; Pred. No. 4.4e+03;

Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ctatagctaggtatctgtc 20
||||| | ||| ||| ||| |||
Db 22 CTCGAGGCTAGTCTCTCTC 3

RESULT 8

US-09-866-108-2604

; Sequence 2604, Application US/09866108

; GENERAL INFORMATION:

; APPLICANT: GU, Yizhong

; APPLICANT: JI, Yonggang

; APPLICANT: PENN, Sharron G.

; APPLICANT: HANZEL, David K.

; APPLICANT: RANK, David R.

; APPLICANT: CHEN, Wensheng

; APPLICANT: SHANNON, Mark

; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE

; FILE REFERENCE: AEOMICA-7

; CURRENT APPLICATION NUMBER: US/09/866,108

; CURRENT FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 60/266,860

; PRIOR FILING DATE: 2001-02-05

; NUMBER OF SEQ ID NOS: 15752

; SOFTWARE: Aecomica Sequence Listing Engine

; SEQ ID NO 2604

; LENGTH: 17

; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-866-108-2604

Query Match 59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 86.7%; Pred. No. 4.9e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agatgctaggatct 17

||||| |||||

Db 3 agatgctggcatct 17

RESULT 9

US-09-866-108-2605

; Sequence 2605, Application US/09866108

; GENERAL INFORMATION:

; APPLICANT: GU, Yizhong

; APPLICANT: JI, Yonggang

; APPLICANT: PENN, Sharron G.

; APPLICANT: HANZEL, David K.

; APPLICANT: RANK, David R.

; APPLICANT: CHEN, Wensheng

; APPLICANT: SHANNON, Mark

; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE

; FILE REFERENCE: AEMICA-7

; CURRENT APPLICATION NUMBER: US/09/866,108

; CURRENT FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 60/266,860

; PRIOR FILING DATE: 2001-02-05

; NUMBER OF SEQ ID NOS: 15752

; SOFTWARE: Aemica Sequence Listing Engine

; SEQ ID NO 2605

; LENGTH: 17

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-866-108-2605

Query Match 59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 86.7%; Pred. No. 4.9e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agatgctaggatct 17

||||| |||||

Db 2 agatgctggcatct 16

RESULT 10

US-09-866-108-2606

; Sequence 2606, Application US/09866108

; GENERAL INFORMATION:

; APPLICANT: GU, Yizhong

; APPLICANT: JI, Yonggang

; APPLICANT: PENN, Sharron G.

; APPLICANT: HANZEL, David K.

; APPLICANT: RANK, David R.

; APPLICANT: CHEN, Wensheng

; APPLICANT: SHANNON, Mark

; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE

; FILE REFERENCE: AEMICA-7

; CURRENT APPLICATION NUMBER: US/09/866,108

; CURRENT FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 60/266,860

; PRIOR FILING DATE: 2001-02-05

; NUMBER OF SEQ ID NOS: 15752

; SOFTWARE: Aemica Sequence Listing Engine

; SEQ ID NO 2606

; LENGTH: 17

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-866-108-2606

Query Match 59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 86.7%; Pred. No. 4.9e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agatgctaggatct 17

||||| |||||

Db 1 agatgctggcatct 15

RESULT 11

US-09-866-108-5533

; Sequence 5533, Application US/09866108

; GENERAL INFORMATION:

; APPLICANT: GU, Yizhong

; APPLICANT: JI, Yonggang

; APPLICANT: PENN, Sharron G.

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 5535
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-5535

Query Match 59.0%; Score 11.8; DB 6; Length 25;
Best Local Similarity 86.7%; Pred. No. 5.3e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 agatgctaggtatct 17
||||| |||||
Db 9 agatgctggcatct 23

RESULT 14

US-09-866-108-5536
; Sequence 5536, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 5536
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-5536

Query Match 59.0%; Score 11.8; DB 6; Length 25;
Best Local Similarity 86.7%; Pred. No. 5.3e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 agatgctaggtatct 17
||||| |||||
Db 8 agatgctggcatct 22

RESULT 15

US-09-866-108-5537
; Sequence 5537, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aecomica Sequence Listing Engine

; SEQ ID NO 5537
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-5537

Query Match 59.0%; Score 11.8; DB 6; Length 25;
Best Local Similarity 86.7%; Pred. No. 5.3e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agatgctaggtatct 17
||||| || |||
Db 7 agatgctgggcattct 21

Search completed: October 2, 2001, 16:55:01
Job time: 17664 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 21:50:09 ; Search time 17695.9 Seconds
(without alignments)
16.701 Million cell updates/sec

Title: US-09-757-100B-7

Perfect score: 20

Sequence: 1 ttctcccttcggtattctt 20

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 14155048 seqs, 7388405095 residues

Total number of hits satisfying chosen parameters: 4519004

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*

2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*

3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*

4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*

5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*

6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*

7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*

8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*

9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*

10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*

11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*

12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*

13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*

14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*

15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*

16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*

17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*

18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*

19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*

20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*

21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*

22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*

23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*

24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*

25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*

26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*

27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*

28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*

30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*

31: /cgn2_6/ptodata/1/pna/US098_COMB.seq.*

32: /cgn2_6/ptodata/1/pna/US099_COMB.seq.*

33: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*

34: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*

35: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*

36: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*

37: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*

38: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*

39: /cgn2_6/ptodata/1/pna/US099G_COMB.seq.*

40: /cgn2_6/ptodata/1/pna/US099H_COMB.seq.*

41: /cgn2_6/ptodata/1/pna/US099I_COMB.seq.*

42: /cgn2_6/ptodata/1/pna/US099J_COMB.seq.*

43: /cgn2_6/ptodata/1/pna/US099K_COMB.seq.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	100.0	20	1	PCT-US00-18999-7
2	20	100.0	20	17	US-09-377-310-7
3	20	100.0	20	29	US-09-757-100B-7
C	15.4	77.0	44	17	US-09-310-298-486
5	15.2	76.0	36	18	US-09-403-752A-115
6	15	75.0	15	1	PCT-US00-18999-27
7	15	75.0	15	17	US-09-377-310-27
8	15	75.0	15	29	US-09-757-100B-27
9	14.8	74.0	25	55	US-60-233-166-23007
10	14.8	74.0	25	55	US-60-234-017-131414
11	14.2	71.0	25	55	US-60-233-166-280120
12	13.8	69.0	20	25	US-09-657-481A-39
C	13.8	69.0	25	55	US-60-233-166-108604
14	13.8	69.0	25	55	US-60-233-166-167371
15	13.8	69.0	25	55	US-60-234-017-425796
C	13.8	69.0	25	55	US-60-234-017-425802
16	13.8	69.0	33	14	US-09-016-061-83
17	13.8	69.0	33	17	US-09-339-922A-83
18	13.8	69.0	33	17	US-09-339-922A-83
19	13.6	68.0	25	55	US-60-232-638-106175
20	13.6	68.0	25	55	US-60-233-166-168611
C	13.6	68.0	25	55	US-60-233-166-330256
22	13.6	68.0	50	12	US-08-899-858-2789
C	13.6	68.0	50	34	US-60-023-277-2789
24	13.4	67.0	25	55	US-60-233-620-25074
25	13.2	66.0	18	53	US-60-216-745-6364
C	13.2	66.0	20	18	US-09-422-978-8164
27	13.2	66.0	21	53	US-60-216-745-5644
28	13.2	66.0	24	1	PCT-US00-24784-163
29	13.2	66.0	25	26	US-09-660-220-87881
30	13.2	66.0	25	55	US-60-232-638-66944
31	13.2	66.0	25	55	US-60-232-638-70508
C	13.2	66.0	25	55	US-60-232-638-85391
33	13.2	66.0	25	55	US-60-233-166-205876
34	13.2	66.0	25	55	US-60-233-166-205876
35	13.2	66.0	25	55	US-60-233-166-236573
36	13.2	66.0	25	55	US-60-233-166-333957
37	13.2	66.0	25	55	US-60-233-166-420026
C	13.2	66.0	25	55	US-60-233-166-420026
39	13.2	66.0	25	55	US-60-234-017-43611
C	13.2	66.0	25	55	US-60-234-017-43611
41	13.2	66.0	25	55	US-60-234-017-287051
C	13.2	66.0	25	55	US-60-234-017-297294
43	13.2	66.0	25	55	US-60-234-017-330691
44	13.2	66.0	25	55	US-60-234-017-528222
C	13.2	66.0	25	55	US-60-234-049-129873

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*

45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*

46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*

47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*

48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*

49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*

50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*

51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*

52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*

53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*

54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*

55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*

56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*

57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*

58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*

59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*

60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*

ALIGNMENTS

RESULT 1
PCT-US00-18999-7
; Sequence 7, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-7

Query Match 100.0%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
| | | | | | | | | | | | | | | | | |
Db 1 ttctcccttcggtattctt 20

RESULT 2
US-09-377-310-7
; Sequence 7, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-7

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
| | | | | | | | | | | | | | | | | |
Db 1 ttctcccttcggtattctt 20

RESULT 3
US-09-757-100b-7
; Sequence 7, Application US/09757100B

; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100b-7

Query Match 100.0%; Score 20; DB 29; Length 20;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
| | | | | | | | | | | | | | | | | |
Db 1 ttctcccttcggtattctt 20

RESULT 4
US-09-310-298-486/c
; Sequence 486, Application US/09310298
; GENERAL INFORMATION:
; APPLICANT: Lander, David G.
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: PRIMER SEQUENCES FOR SINGLE NUCLEOTIDE
; FILE REFERENCE: WHIFG98-07pA
; CURRENT APPLICATION NUMBER: US/09/310,298
; CURRENT FILING DATE: 1998-05-12
; EARLIER APPLICATION NUMBER: 60/085,152
; PRIOR FILING DATE: 1998-05-12
; NUMBER OF SEQ ID NOS: 5968
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 486
; LENGTH: 44
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide primer
US-09-310-298-486

Query Match 77.0%; Score 15.4; DB 17; Length 44;
Best Local Similarity 94.1%; Pred. No. 2.7e+03;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 tcccttcggtattctt 20
| | | | | | | | | | | | | | | | | |
Db 39 TCCCTTCGTTATTCAT 23

RESULT 5
US-09-403-752A-115
; Sequence 115, Application US/09403752A
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: RICIN-LIKE TOXIN VARIANTS FOR TREATMENT OF CANCER,
; OTHER INFORMATION: VIRAL OR PARASITIC INFECTIONS


```
; FILE REFERENCE: 10447-005
; CURRENT APPLICATION NUMBER: US/09/403,752A
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: U.S. 60/045,148
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: U.S. 60/063,715
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 115
; LENGTH: 36
; TYPE: DNA
; ORGANISM: Human Prostate-Specific Antigen linker region of PAP-290
US-09-403-752A-115

Query Match          76.0%; Score 15.2; DB 18; Length 36;
Best Local Similarity 85.0%; Pred. No. 3.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctccctccgttattctt 20
Db 14 ttctcttccgatattttt 33

RESULT 6
PCT-US00-18999-27
; Sequence 27, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-27

Query Match          75.0%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ctccctccgttatt 17
Db 1 ctccctccgttatt 15

RESULT 7
US-09-377-310-27
; Sequence 27, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0

; FILE REFERENCE: 10447-005
; CURRENT APPLICATION NUMBER: US/09/403,752A
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: U.S. 60/045,148
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: U.S. 60/063,715
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-27

Query Match          75.0%; Score 15; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ctccctccgttatt 17
Db 1 ctccctccgttatt 15

RESULT 8
US-09-757-100B-27
; Sequence 27, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-27

Query Match          75.0%; Score 15; DB 29; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ctccctccgttatt 17
Db 1 ctccctccgttatt 15

RESULT 9
US-60-233-166-23007
; Sequence 23007, Application US/60233166
; GENERAL INFORMATION:
; APPLICANT: Mitmann
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 3112
; CURRENT APPLICATION NUMBER: US/60/233,166
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23007
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AA818353
```

US-60-233-166-23007

Query Match 74.0%; Score 14.8; DB 55; Length 25;
Best Local Similarity 88.9%; Pred. No. 4.8e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 ctccctccggtattctt 20
||||| ||| |||

Db 8 ctccctccggtattctt 25

RESULT 10

US-60-234-017-131414

; Sequence 131414, Application US/60234017

; GENERAL INFORMATION:

; APPLICANT: Mittmann, M

; APPLICANT: Affymetrix, Inc.

; TITLE OF INVENTION: Methods of Genetic Analysis of Mus

; FILE REFERENCE: 3115

; CURRENT APPLICATION NUMBER: US/60/234,017

; CURRENT FILING DATE: 2000-09-20

; NUMBER OF SEQ ID NOS: 605887

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 131414

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Mus musculus

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: GenBank AA600645

US-60-234-017-131414

Query Match

74.0%; Score 14.8; DB 55; Length 25;
Best Local Similarity 88.9%; Pred. No. 4.8e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 tctccctccggtattctt 19
||| ||| ||| ||| ||| |||

Db 8 tcaccttccggtattctt 25

RESULT 11

US-60-233-166-280120

; Sequence 280120, Application US/60233166

; GENERAL INFORMATION:

; APPLICANT: Mittmann

; APPLICANT: Affymetrix, Inc.

; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat

; FILE REFERENCE: 3112

; CURRENT APPLICATION NUMBER: US/60/233,166

; CURRENT FILING DATE: 2000-10-24

; NUMBER OF SEQ ID NOS: 420907

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 280120

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: GenBank AF639138

US-60-233-166-280120

Query Match

71.0%; Score 14.2; DB 55; Length 25;
Best Local Similarity 84.2%; Pred. No. 9.2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ttctccctccggtattctt 19
||| ||| ||| ||| ||| |||

Db 1 ttccgacctccggtattctt 19

RESULT 12

US-09-657-481A-39/c

; Sequence 39, Application US/09657481A

; GENERAL INFORMATION:

; APPLICANT: Brett P. Monia

; APPLICANT: Lex M. Cowser

; TITLE OF INVENTION: ANTISENSE MODULATION OF UBIQUITIN PROTEIN LIGASE WWP1 AND

; FILE REFERENCE: ETS-0087

; CURRENT APPLICATION NUMBER: US/09/657,481A

; CURRENT FILING DATE: 2000-09-07

; NUMBER OF SEQ ID NOS: 93

; SEQ ID NO 39

; LENGTH: 20

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Antisense Oligonucleotide

US-09-657-481A-39

Query Match

69.0%; Score 13.8; DB 25; Length 20;
Best Local Similarity 88.2%; Pred. No. 1.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ttctccctccggtattctt 17

||| ||| ||| ||| ||| |||

Db 18 TGCTCACTCCGTTATT 2

RESULT 13

US-60-233-166-108604/c

; Sequence 108604, Application US/60233166

; GENERAL INFORMATION:

; APPLICANT: Mittmann

; APPLICANT: Affymetrix, Inc.

; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat

; FILE REFERENCE: 3112

; CURRENT APPLICATION NUMBER: US/60/233,166

; CURRENT FILING DATE: 2000-10-24

; NUMBER OF SEQ ID NOS: 420907

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 108604

; LENGTH: 25

; TYPE: DNA

; ORGANISM: Rattus norvegicus

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: GenBank AA860043

US-60-233-166-108604

Query Match

69.0%; Score 13.8; DB 55; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 3 ctccctccggtattctt 19

||||| ||| ||| ||| |||

Db 21 CTCCTACCGTTGTCT 5

RESULT 14

US-60-233-166-167371

; Sequence 167371, Application US/60233166

; GENERAL INFORMATION:

; APPLICANT: Mittmann

; APPLICANT: Affymetrix, Inc.

; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat

; FILE REFERENCE: 3112

; CURRENT APPLICATION NUMBER: US/60/233,166

; CURRENT FILING DATE: 2000-10-24

; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167371
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AA926247
US-60-233-166-167371

Query Match 69.0%; Score 13.8; DB 55; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 tcccttcggttattctt 20
||||| ||||| ||||
Db 5 tcccttcggttattctt 21

RESULT 15
US-60-234-017-425796/c
; Sequence 425796, Application US/60234017
; GENERAL INFORMATION:
; APPLICANT: Mittmann, M
; TITLE OF INVENTION: Methods of Genetic Analysis of Mus
; TITLE OF INVENTION: musculus
; FILE REFERENCE: 3115
; CURRENT APPLICATION NUMBER: US/60/234,017
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 605887
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 425796
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AV102460
US-60-234-017-425796

Query Match 69.0%; Score 13.8; DB 55; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 tctcccttcggttatttc 18
||||| ||||| ||||
Db 18 TCTCCCTTCGCTAATC 2

Search completed: October 2, 2001, 21:50:10
Job time: 24523 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:55:00 ; Search time 2173.58 Seconds
(without alignments)
19.290 Million cell updates/sec

Title: US-09-757-100B-7

Perfect score: 20

Sequence: 1 ttctcccttcggtattctt 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1895189 seqs, 1048201267 residues

Total number of hits satisfying chosen parameters: 299022

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	15.2	76.0	36	7	US-09-551-151-115
2	13.8	69.0	21	6	US-09-726-774-70
3	13.8	69.0	33	6	US-09-900-590-83
4	13.6	68.0	36	8	US-60-252-833-32245
5	12.6	63.0	30	4	US-08-850-062A-75
6	12.6	63.0	46	8	US-60-278-561-657
7	12.2	61.0	47	6	US-09-516-667-73
8	12.2	61.0	48	6	US-09-516-667-35
9	12.2	61.0	48	6	US-09-516-667-36
10	12.2	61.0	48	6	US-09-516-667-74
11	12.2	61.0	48	6	US-09-516-667-75
12	12.2	60.0	45	8	US-60-252-833-9824
13	11.8	59.0	28	8	US-60-253-652-14008
14	11.6	58.0	29	8	US-60-253-456-15489
15	11.6	58.0	31	5	US-09-574-376B-90
16	11.6	58.0	31	5	US-09-574-376B-1240
17	11.6	58.0	37	5	US-09-708-690-18891
18	11.6	58.0	38	5	US-09-532-537B-1781
19	11.6	58.0	46	6	US-09-845-674-616
20	11.4	57.0	20	5	US-09-632-703-22
21	11.4	57.0	20	6	US-09-198-452A-1772
22	11.4	57.0	20	6	US-09-828-707-7
23	11.2	56.0	17	5	US-09-371-772B-5150
24	11.2	56.0	17	5	US-09-708-690-5150
25	11.2	56.0	19	7	US-09-747-391-16

26	11.2	56.0	20	6	US-09-198-452A-6351	Sequence 6351, Ap
27	11.2	56.0	25	6	US-09-881-565-19	Sequence 19, Appl
28	11.2	56.0	28	6	US-09-343-126B-50	Sequence 50, Appl
29	11.2	56.0	28	6	US-09-343-126B-51	Sequence 51, Appl
30	11.2	56.0	28	6	US-09-488-265-63	Sequence 63, Appl
31	11.2	56.0	31	6	US-09-801-274-736	Sequence 736, App
32	11.2	56.0	31	6	US-09-801-274-793	Sequence 793, App
33	11.2	56.0	31	6	US-09-801-274-1470	Sequence 1470, Ap
34	11.2	56.0	38	6	US-09-546-745A-5238	Sequence 5238, Ap
35	11.2	56.0	43	8	US-60-253-653-10876	Sequence 10876, A
36	11.2	56.0	50	5	US-09-699-011A-293	Sequence 293, App
37	11.2	56.0	50	6	US-09-881-565-22	Sequence 22, Appl
38	11.2	55.0	21	6	US-09-765-081-247	Sequence 247, App
39	11.2	55.0	23	7	US-09-851-271A-13	Sequence 13, Appl
40	11.2	55.0	28	7	US-09-735-271-1880	Sequence 1880, Ap
41	11.2	55.0	37	5	US-09-708-690-18282	Sequence 18282, A
42	11.2	55.0	37	5	US-09-708-690-18422	Sequence 18422, A
43	11.2	55.0	37	5	US-09-708-690-18674	Sequence 18674, A
44	11.2	55.0	37	5	US-09-708-690-18740	Sequence 18740, A
45	11.2	55.0	37	5	US-09-708-690-20466	Sequence 20466, A

ALIGNMENTS

RESULT 1

US-09-551-151-115
; Sequence 115, Application US/095511151
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; FILE REFERENCE: 10447-011
; CURRENT APPLICATION NUMBER: US/09/551,151
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 115
; LENGTH: 36
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human Prostate-Specific Antigen linker region of pAP-290
US-09-551-151-115

Query Match 76.0%; Score 15.2; DB 7; Length 36;
Best Local Similarity 85.0%; Pred. No. 4.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattctt 20

Db 14 ttctctcttcggtattttt 33

RESULT 2

US-09-726-774-70
; Sequence 70, Application US/09726774
; GENERAL INFORMATION:
; APPLICANT: Iversen, Patrick L.
; TITLE OF INVENTION: Antisense Antibacterial Method and
; FILE REFERENCE: 0450-0032.30
; CURRENT APPLICATION NUMBER: US/09/726,774
; CURRENT FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: US 60/168,150
; PRIOR FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 21

```
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense oligomer
US-09-726-774-70

Query Match      69.0%; Score 13.8; DB 6; Length 21;
Best Local Similarity 88.2%; Pred. No. 2.1e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ttctccctccggtattc 18
   | ||||| |||||
Db 3 ttctccctccgtaattc 19

RESULT 3
US-09-900-590-83
; Sequence 83, Application US/09900590
; GENERAL INFORMATION:
; APPLICANT: Huse, William D.
; TITLE OF INVENTION: Anti-alpha V Beta 3 Recombinant Human
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,590
; FILING DATE: 06-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/016,061
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-IX 2965
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..33
; SEQUENCE DESCRIPTION: SEQ ID NO: 83:
US-09-900-590-83

Query Match      69.0%; Score 13.8; DB 6; Length 33;
Best Local Similarity 88.2%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 ttctccctccggtattc 20
   | | | | | | | | | |
Db 3 TCTCATCCGTTATTCTT 19
```

```
RESULT 4
US-60-252-833-32245/c
; Sequence 32245, Application US/60252833
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R
; TITLE OF INVENTION: Compositions isolated from bovine
; FILE REFERENCE: 1052P2
; CURRENT APPLICATION NUMBER: US/60/252,833
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 43535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32245
; LENGTH: 36
; TYPE: DNA
; ORGANISM: Bovine
US-60-252-833-32245

Query Match      68.0%; Score 13.6; DB 8; Length 36;
Best Local Similarity 80.0%; Pred. No. 2.8e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttctccctccggtattc 20
   | | | | | | | | | |
Db 34 TTCTCCCTTGCTAGTTCTT 15

RESULT 5
US-08-850-062A-75
; Sequence 75, Application US/08850062A
; GENERAL INFORMATION:
; APPLICANT: Skolnick, Mark H.
; Goldgar, David E.
; Miki, Yoshio
; Swenson, Jeff
; Kamb, Alexander
; Harshman, Keith D.
; Shattuck-Eidens, Donna M.
; Tavtigian, Sean V.
; Wiseman, Roger W.
; Futreal, P. Andrew
; TITLE OF INVENTION: 17q-Linked Breast and Ovarian Cancer
; Susceptibility Gene
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Manbeck, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701-E
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,062A
; FILING DATE: 02-May-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/483,554
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: US 08/409,305
; FILING DATE: 24-MAR-1995
; APPLICATION NUMBER: US 08/348,824
; FILING DATE: 29-NOV-1994
; APPLICATION NUMBER: US 08/308,104
; FILING DATE: 16-SEP-1994
; APPLICATION NUMBER: US 08/300,266
; FILING DATE: 02-SEP-1994
```

; APPLICATION NUMBER: US 08/289,221
; FILING DATE: 12-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ihnen, Jeffrey L.
; REGISTRATION NUMBER: 28,957
; REFERENCE/DOCKET NUMBER: 2318-137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-783-6040
; TELEFAX: 202-9783-6031
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-08-850-062A-75

Query Match 63.0%; Score 12.6; DB 4; Length 30;
Best Local Similarity 78.9%; Pred. No. 8.1e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ttctcccttcggtattctt 20
||||| ||||| ||||| |||||
DB 8 TCTCTCTTCCTCTCTCT 26

RESULT 6
US-60-278-561-657/c
; Sequence 657, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; Nucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 657
; LENGTH: 46
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: SNP00057293
; NAME/KEY: snp
; LOCATION: 21
; OTHER INFORMATION: 017559.1, 21, G->A
US-60-278-561-657

Query Match 63.0%; Score 12.6; DB 8; Length 46;
Best Local Similarity 78.9%; Pred. No. 8.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattctt 19
||||| ||||| ||||| |||||
DB 22 TTCCTCTTCCTCTCTCT 4

RESULT 7
US-09-516-667-73/c
; Sequence 73, Application US/09516667
; GENERAL INFORMATION:

; APPLICANT: Inouye, Masayori
; APPLICANT: Wang, Nan
; APPLICANT: Yamanaka, Kunitoshi
; TITLE OF INVENTION: COLD-SHOCK REGULATORY ELEMENTS, CONSTRUCTS THEREOF, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 1053-00
; CURRENT APPLICATION NUMBER: US/09/516,667
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 47
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-09-516-667-73

Query Match 61.0%; Score 12.2; DB 6; Length 47;
Best Local Similarity 82.4%; Pred. No. 1.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattt 17
||||| ||||| |||||
DB 30 TTCCTCTTCCTCTCTT 14

RESULT 8
US-09-516-667-35/c
; Sequence 35, Application US/09516667
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Wang, Nan
; APPLICANT: Yamanaka, Kunitoshi
; TITLE OF INVENTION: COLD-SHOCK REGULATORY ELEMENTS, CONSTRUCTS THEREOF, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 1053-00
; CURRENT APPLICATION NUMBER: US/09/516,667
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-09-516-667-35

Query Match 61.0%; Score 12.2; DB 6; Length 48;
Best Local Similarity 82.4%; Pred. No. 1.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattt 17
||||| ||||| ||||| |||||
DB 29 TTCCTCTTCCTCTTATT 13

RESULT 9
US-09-516-667-36
; Sequence 36, Application US/09516667
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Wang, Nan
; APPLICANT: Yamanaka, Kunitoshi
; TITLE OF INVENTION: COLD-SHOCK REGULATORY ELEMENTS, CONSTRUCTS THEREOF, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 1053-00
; CURRENT APPLICATION NUMBER: US/09/516,667

; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-09-516-667-36

Query Match 61.0%; Score 12.2; DB 6; Length 48;
Best Local Similarity 82.4%; Pred. No. 1.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctccctccggtatt 17
||| ||||| |||||
DB 24 ttcccccttcattatt 40

RESULT 10
US-09-516-667-74/c
; Sequence 74, Application US/09516667
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Wang, Nan
; APPLICANT: Yamanaka, Kunitoshi
; TITLE OF INVENTION: COLD-SHOCK REGULATORY ELEMENTS, CONSTRUCTS THEREOF, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 1053-00
; CURRENT APPLICATION NUMBER: US/09/516,667
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 48
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-09-516-667-74

Query Match 61.0%; Score 12.2; DB 6; Length 48;
Best Local Similarity 82.4%; Pred. No. 1.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctccctccggtatt 17
||| ||||| |||||
DB 30 TTCCCCCTTCATTATT 14

RESULT 11
US-09-516-667-75/c
; Sequence 75, Application US/09516667
; GENERAL INFORMATION:
; APPLICANT: Inouye, Masayori
; APPLICANT: Wang, Nan
; APPLICANT: Yamanaka, Kunitoshi
; TITLE OF INVENTION: COLD-SHOCK REGULATORY ELEMENTS, CONSTRUCTS THEREOF, AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 1053-00
; CURRENT APPLICATION NUMBER: US/09/516,667
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 48
; TYPE: RNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: oligonucleotide
US-09-516-667-75

Query Match 61.0%; Score 12.2; DB 6; Length 48;
Best Local Similarity 82.4%; Pred. No. 1.3e+04;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ttctccctccggtatt 17
||| ||||| |||||
DB 30 TTCCCCCTTCATTATT 14

RESULT 12
US-60-252-833-9824
; Sequence 9824, Application US/60252833
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R
; TITLE OF INVENTION: Compositions isolated from bovine
; TITLE OF INVENTION: tissues and methods for their use.
; FILE REFERENCE: 1052P2
; CURRENT APPLICATION NUMBER: US/60/252,833
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 43535
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9824
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Bovine
US-60-252-833-9824

Query Match 60.0%; Score 12; DB 8; Length 45;
Best Local Similarity 75.0%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ttctccctccggtatttt 20
||| ||||| |||||
DB 3 tctgcctctgtgaagtctt 22

RESULT 13
US-60-253-652-14008
; Sequence 14008, Application US/60253652
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R
; TITLE OF INVENTION: Compositions isolated from bovine
; TITLE OF INVENTION: tissues and methods for their use.
; FILE REFERENCE: 1055P2
; CURRENT APPLICATION NUMBER: US/60/253,652
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 29954
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14008
; LENGTH: 28
; TYPE: DNA
; ORGANISM: Bovine
US-60-253-652-14008

Query Match 59.0%; Score 11.8; DB 8; Length 28;
Best Local Similarity 86.7%; Pred. No. 1.9e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ctccctccggtatt 17
| ||||| |||||
DB 1 caccctccggtctt 15

RESULT 14

US-60-253-456-15489
; Sequence 15489, Application US/60253456
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J.
; TITLE OF INVENTION: Polynucleotides, isolated from
; TITLE OF INVENTION: plants, and methods for their use.
; FILE REFERENCE: 1054P1
; CURRENT APPLICATION NUMBER: US/60/253,456
; CURRENT FILING DATE: 2000-11-27
; NUMBER OF SEQ ID NOS: 37096
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15489
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Pinus radiata
US-60-253-456-15489

Query Match 58.0%; Score 11.6; DB 8; Length 29;
Best Local Similarity 77.8%; Pred. No. 2.4e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 tctcccttcggtattctt 19
| | | | | | | | | |
Db 1 ttccctcactggtattct 18

RESULT 15

US-09-574-376B-90/c
; Sequence 90, Application US/09574376B
; GENERAL INFORMATION:
; APPLICANT: Warrington, Janet
; APPLICANT: Shah, Nila
; APPLICANT: Gingeras, Thomas Raymond
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Human Lymphoblast Polymorphisms
; FILE REFERENCE: 3229.2
; CURRENT APPLICATION NUMBER: US/09/574,376B
; CURRENT FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 1330
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 90
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
US-09-574-376B-90

Query Match 58.0%; Score 11.6; DB 5; Length 31;
Best Local Similarity 70.0%; Pred. No. 2.4e+04;
Matches 14; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
| | | | | | | | | |
Db 27 TTGTCACCTTGCTACTTCTT 8

Search completed: October 2, 2001, 16:55:00
Job time: 17663 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 21:50:08 ; Search time 17695.9 Seconds
(without alignments)
16.701 Million cell updates/sec

Title: US-09-757-100B-6
Perfect score: 20
Sequence: 1 gaaactgcagaagcactga 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 14155048 seqs, 738840595 residues
Total number of hits satisfying chosen parameters: 4519004

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US098_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	100.0	20	1	PCT-US00-18999-6
2	20	100.0	20	17	US-09-757-310-6
3	20	100.0	20	29	US-09-757-100B-6
4	15.2	76.0	25	55	US-60-233-166-247462
5	15	75.0	15	1	PCT-US00-18999-26
6	15	75.0	15	17	US-09-757-310-26
7	15	75.0	15	29	US-09-757-100B-26
8	15	75.0	25	55	US-60-233-166-169359
9	15	75.0	25	55	US-60-233-166-169364
10	14.8	74.0	25	26	US-09-660-080-2372
11	14.8	74.0	25	26	US-09-660-220-7628
12	14.8	74.0	25	55	US-60-233-166-75968
13	14.4	72.0	25	55	US-60-233-166-247458
14	14.2	71.0	25	55	US-60-233-166-353412
15	14.2	71.0	25	55	US-60-233-357-7140
16	14.2	71.0	45	17	US-09-310-298-1603
17	13.8	69.0	20	31	US-09-844-521-16
18	13.8	69.0	25	55	US-60-233-166-250484
19	13.8	69.0	25	55	US-60-233-166-250515
20	13.8	69.0	25	55	US-60-233-620-33139
21	13.8	69.0	25	55	US-60-233-620-33145
22	13.8	69.0	25	55	US-60-233-620-33146
23	13.8	69.0	25	55	US-60-234-017-31881
24	13.8	69.0	25	55	US-60-234-017-31890
25	13.8	69.0	25	55	US-60-234-049-10183
26	13.6	68.0	25	55	US-60-233-166-142391
27	13.6	68.0	25	55	US-60-233-166-341643
28	13.6	68.0	25	55	US-60-233-166-358933
29	13.6	68.0	25	55	US-60-233-357-14127
30	13.6	68.0	25	55	US-60-234-017-150021
31	13.6	68.0	25	55	US-60-234-017-427143
32	13.6	68.0	25	55	US-60-234-017-427147
33	13.6	68.0	25	55	US-60-234-049-131450
34	13.6	68.0	42	17	US-09-318-138-15
35	13.4	67.0	19	28	US-09-703-708-12274
36	13.4	67.0	19	48	US-60-164-320-12274
37	13.4	67.0	19	50	US-60-183-791-12274
38	13.4	67.0	20	16	US-09-201-228A-4597
39	13.4	67.0	25	26	US-09-660-080-2373
40	13.4	67.0	25	55	US-60-233-166-75953
41	13.4	67.0	25	55	US-60-233-620-33138
42	13.4	67.0	36	16	US-09-232-278A-25
43	13.4	67.0	43	1	PCT-US96-13455-12
44	13.4	67.0	43	15	US-09-116-115-12
45	13.4	67.0	43	21	US-09-541-762-12

ALIGNMENTS

RESULT 1
PCT-US00-18999-6
; Sequence 6, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-6

Query Match 100.0%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaaactgcagaaggcactga 20
Db 1 gaaactgcagaaggcactga 20
|||||

RESULT 2
US-09-377-310-6
; Sequence 6, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-6

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaaactgcagaaggcactga 20
Db 1 gaaactgcagaaggcactga 20
|||||

RESULT 3
US-09-757-100B-6
; Sequence 6, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-6

; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-6

Query Match 100.0%; Score 20; DB 29; Length 20;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaaactgcagaaggcactga 20
Db 1 gaaactgcagaaggcactga 20
|||||

RESULT 4
US-60-233-166-247462/C
; Sequence 247462, Application US/60233166
; GENERAL INFORMATION:
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat
; FILE REFERENCE: 3112
; CURRENT APPLICATION NUMBER: US/60/233,166
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 247462
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AI043982
US-60-233-166-247462

Query Match 76.0%; Score 15.2; DB 55; Length 25;
Best Local Similarity 85.0%; Pred. No. 3.1e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 gaaactgcagaaggcactga 20
Db 22 GAGTCAGCAGAGGCACTGA 3
|||

RESULT 5
PCT-US00-18999-26
; Sequence 26, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-6

; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-26

Query Match 75.0%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 aactgcagaagggc 17
|||||
DB 1 aactgcagaagggc 15

RESULT 6
US-09-377-310-26
; Sequence 26, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-26

Query Match 75.0%; Score 15; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 aactgcagaagggc 17
|||||
DB 1 aactgcagaagggc 15

RESULT 7
US-09-757-100B-26
; Sequence 26, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13

; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-26

Query Match 75.0%; Score 15; DB 29; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.5e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 aactgcagaagggc 17
|||||
DB 1 aactgcagaagggc 15

RESULT 8
US-60-233-166-169359/c
; Sequence 169359, Application US/60233166
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 3112
; CURRENT APPLICATION NUMBER: US/60/233,166
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169359
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AA955408
US-60-233-166-169359

Query Match 75.0%; Score 15; DB 55; Length 25;
Best Local Similarity 100.0%; Pred. No. 3.9e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaaactgcagaagggc 15
|||||
DB 18 GAAACTGCAGAAGGC 4

RESULT 9
US-60-233-166-169364/c
; Sequence 169364, Application US/60233166
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat
; TITLE OF INVENTION: Genome
; FILE REFERENCE: 3112
; CURRENT APPLICATION NUMBER: US/60/233,166
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169364
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AA955408
US-60-233-166-169364

; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 353412
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank U59126
US-60-233-166-353412

Query Match 71.0%; Score 14.2; DB 55; Length 25;
Best Local Similarity 84.2%; Pred. No. 9.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 aaactgcagaaggcactga 20
||||| ||||| ||| |
Db 2 aaactacagaaggtactca 20

RESULT 15
US-60-233-357-7140
; Sequence 7140, Application US/60233357
; GENERAL INFORMATION:
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of Rat
; FILE REFERENCE: 3114
; CURRENT APPLICATION NUMBER: US/60/233,357
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 21305
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7140
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank D49836
US-60-233-357-7140

Query Match 71.0%; Score 14.2; DB 55; Length 25;
Best Local Similarity 84.2%; Pred. No. 9.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 aaactgcagaaggcactga 20
||||| ||||| ||| |
Db 2 aaactacagaaggtactca 20

Search completed: October 2, 2001, 21:50:09
Job time: 24522 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:54:59 ; Search time 2173.58 Seconds
(without alignments)
19.290 Million cell updates/sec

Title: US-09-757-100B-6

Perfect score: 20

Sequence: 1 gaaactgcagaagcactga 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1895189 seqs, 1048201267 residues

Total number of hits satisfying chosen parameters: 299022

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Pending_Patents_NA_New.*
- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
 - 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
 - 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
 - 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
 - 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
 - 6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq1.*
 - 7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2.*
 - 8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	14.4	72.0	17	6	US-09-866-108-2022
c 2	14.4	72.0	17	6	US-09-866-108-2023
c 3	14.4	72.0	25	6	US-09-866-108-4951
c 4	14.4	72.0	25	6	US-09-866-108-4952
c 5	14.4	72.0	25	6	US-09-866-108-4953
c 6	14.4	72.0	25	6	US-09-866-108-4954
c 7	14.4	72.0	25	6	US-09-866-108-4955
c 8	14.4	72.0	25	6	US-09-866-108-4956
c 9	14.4	72.0	25	6	US-09-866-108-4957
c 10	14.4	72.0	25	6	US-09-866-108-4958
c 11	14.4	72.0	25	6	US-09-866-108-4959
c 12	14.4	72.0	25	6	US-09-866-108-4960
c 13	13.8	69.0	31	6	US-09-801-274-1595
c 14	13.4	67.0	17	6	US-09-866-108-2021
c 15	13.4	67.0	17	6	US-09-866-108-2024
c 16	13.4	67.0	25	6	US-09-866-108-4950
c 17	13.4	67.0	25	6	US-09-866-108-4961
c 18	13	65.0	17	6	US-09-866-108-2025
c 19	13	65.0	17	6	US-09-866-108-2026
c 20	13	65.0	25	6	US-09-866-108-4962
c 21	13	65.0	25	6	US-09-866-108-4963
c 22	12.6	63.0	24	7	US-09-919-478-18
c 23	12.6	63.0	25	7	US-09-724-648-6
c 24	12.6	63.0	26	6	US-09-902-214-52
c 25	12.6	63.0	46	7	US-09-919-478-15

ALIGNMENTS

RESULT 1

US-09-866-108-2022/c

; Sequence 2022, Application US/09866108

; GENERAL INFORMATION:

; APPLICANT: GU, Yizhong

; APPLICANT: JI, Yonggang

; APPLICANT: PENN, Sharron G.

; APPLICANT: HANZEL, David K.

; APPLICANT: RANK, David R.

; APPLICANT: CHEN, Wensheng

; TITLE OF INVENTION: SHANNON, Mark

; FILE REFERENCE: AEGMICA-7

; CURRENT APPLICATION NUMBER: US/09/866.108

; CURRENT FILING DATE: 2001-05-25

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: PCT/US01/00666

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00661

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 60/266,860

; PRIOR FILING DATE: 2001-02-05

; NUMBER OF SEQ ID NOS: 15752

; SOFTWARE: AEGMICA Sequence Listing Engine

; SEQ ID NO 2022

Sequence 2020, Ap
Sequence 4949, Ap
Sequence 7, Appli
Sequence 33, Appli
Sequence 223, App
Sequence 3, Appli
Sequence 2, Appli
Sequence 2027, Ap
Sequence 6556, Ap
Sequence 4964, Ap
Sequence 45, Appli
Sequence 1777, Ap
Sequence 370, App
Sequence 370, App
Sequence 11980, A
Sequence 14257, A
Sequence 18382, A
Sequence 2, Appli
Sequence 39, Appli


```
; APPLICANT: JL, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 4952
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-4952

Query Match 72.0%; Score 14.4; DB 6; Length 25;
Best Local Similarity 93.8%; Pred. No. 1.le+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ctgcagaagcactga 20
Db 24 CTGCAGAAGGCACCGA 9

RESULT 5
US-09-866-108-4953/c
; Sequence 4953, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JL, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
```

```
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 4953
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-4953

Query Match 72.0%; Score 14.4; DB 6; Length 25;
Best Local Similarity 93.8%; Pred. No. 1.le+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ctgcagaagcactga 20
Db 23 CTGCAGAAGGCACCGA 8

RESULT 6
US-09-866-108-4954/c
; Sequence 4954, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JL, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
```

; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 4954
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-4954

Query Match 72.0%; Score 14.4; DB 6; Length 25;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ctgcagaagcactga 20
|||||
DB 22 CTGCAGAAGGCACCGA 7

RESULT 7
US-09-866-108-4955/c
; Sequence 4955, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 15752
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 4955
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-108-4955

Query Match 72.0%; Score 14.4; DB 6; Length 25;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ctgcagaagcactga 20
|||||
DB 21 CTGCAGAAGGCACCGA 6

RESULT 8
US-09-866-108-4956/c
; Sequence 4956, Application US/09866108
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharron G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AEOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/266,860
; PRIOR FILING DATE: 2001-02-05

Sequence 4959, Application US/09866108
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharon G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
FILE REFERENCE: AEOMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/266,860
PRIOR FILING DATE: 2001-02-05
NUMBER OF SEQ ID NOS: 15752
SOFTWARE: Aeomica Sequence Listing Engine
SEQ ID NO 4959
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapiens
US-09-866-108-4959

Query Match 72.0%; Score 14.4; DB 6; Length 25;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ctgcagaaggcactga 20
|||||
Db 17 CTGCAGAGGCACCGA 2

RESULT 12
US-09-866-108-4960/c
Sequence 4960, Application US/09866108
GENERAL INFORMATION:
APPLICANT: GU, Yizhong
APPLICANT: JI, Yonggang
APPLICANT: PENN, Sharon G.
APPLICANT: HANZEL, David K.
APPLICANT: RANK, David R.
APPLICANT: CHEN, Wensheng
APPLICANT: SHANNON, Mark
TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE

FILE REFERENCE: AEOMICA-7
CURRENT APPLICATION NUMBER: US/09/866,108
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/266,860
PRIOR FILING DATE: 2001-02-05
NUMBER OF SEQ ID NOS: 15752
SOFTWARE: Aeomica Sequence Listing Engine
SEQ ID NO 4960
LENGTH: 25
TYPE: DNA
ORGANISM: Homo sapiens
US-09-866-108-4960

Query Match 72.0%; Score 14.4; DB 6; Length 25;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ctgcagaaggcactga 20
|||||
Db 16 CTGCAGAGGCACCGA 1

RESULT 13
US-09-801-274-1595
Sequence 1595, Application US/09801274
GENERAL INFORMATION:
APPLICANT: Cargill, Michele
APPLICANT: Ireland, James S.
APPLICANT: Lander, Eric S.
TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
FILE REFERENCE: 2825.2009-001
CURRENT APPLICATION NUMBER: US/09/801,274
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 60/187,510
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 60/206,129
PRIOR FILING DATE: 2000-05-22
NUMBER OF SEQ ID NOS: 1802
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1595
LENGTH: 31
TYPE: DNA
ORGANISM: Homo sapiens
US-09-801-274-1595

Query Match	69.0%;	Score 13.8;	DB 6;	Length 31;
Best Local Similarity	78.9%;	Pred. No. 2.3e+03;		
Matches 15;	Conservative 1;	Mismatches 3;	Indels	

Qy 1 gaaactgcagaaggcactg 19
||| ||||| |||: |||
Db 2 gacactgcaggaggmgctg 20

RESULT 14

US-09-866--108-2021/c

Sequence 2021, Application US/09866108

GENERAL INFORMATION:

APPLICANT: GU, Yizhong

APPLICANT: JI, Yonggang

APPLICANT: PENN, Sharron G.

APPLICANT: HANZEL, David K.

APPLICANT: RANK, David R.

APPLICANT: CHEN, Wensheng

APPLICANT: SHANNON, Mark

FILE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE

FILE REFERENCE: AEMICA-7

CURRENT APPLICATION NUMBER: US/09/866.108

CURRENT FILING DATE: 2001-05-25

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/006666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/006670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 60/266,860

PRIOR FILING DATE: 2001-02-05

NUMBER OF SEQ ID NOS: 15752

SOFTWARE: Aemica Sequence Listing Engine

SEQ ID NO 2021

LENGTH: 17

TYPE: DNA

ORGANISM: Homo sapiens

US-09-866-108-2021

Query Match 67.0%; Score 13.4; DB 6; Length 17;
Best Local Similarity 93.3%; pred. No. 3.2e+03;
Matches 14; Conservative 0; Mismatches 1; Indels

Qy	6	tgcagaaggcactga	20
Db	17	TGCAGAAGGCACCGA	3

RESULT 15
 US-09-866-108-2024/c
 ; Sequence 2024, Application US/09866108
 ; GENERAL INFORMATION:
 ; APPLICANT: GU, Yizhong
 ; APPLICANT: JI, Yonggang
 ; APPLICANT: PENN, Sharron G.
 ; APPLICANT: HANZEL, David K.
 ; APPLICANT: RANK, David R.
 ; APPLICANT: CHEN, Wensheng
 ; APPLICANT: SHANNON, Mark
 ; TITLE OF INVENTION: MYOSTIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
 ; FILE REFERENCE: AEOMICA-7
 ; CURRENT APPLICATION NUMBER: US/09/866,108
 ; CURRENT FILING DATE: 2001-05-25
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 60/266,860
 ; PRIOR FILING DATE: 2001-02-05
 ; NUMBER OF SEQ ID NOS: 15752
 ; SOFTWARE: Aeomica Sequence Listing Engine
 ; SEQ ID NO 2024
 ; LENGTH: 17
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-866-108-2024

Query Match	67.0%	Score 13.4;	DB 6;	Length 17;
Best Local Similarity	93.3%;	Pred. No. 3.2e+03;		
Matches 14;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps	0;			
Qy	5	ctgcgagaaggcactg	19	
Db	15	CTGCAGAAAGGCACCG	1	

Search completed: October 2, 2001, 16:55:00
Job time: 17663 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:54:58 ; Search time 2173.58 Seconds
(without alignments)
19.290 Million cell updates/sec

Title: US-09-757-100b-4

Perfect score: 20

Sequence: 1 ggcgcgtgaagcgaaggca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1895189 seqs, 1048201267 residues

Total number of hits satisfying chosen parameters: 299022

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq1.*
- 7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2.*
- 8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12.2	61.0	43	8	US-60-252-833-26301
2	12	60.0	35	7	US-09-475-947A-204
3	12	60.0	50	6	US-09-442-849B-1086
4	12	60.0	50	7	US-09-443-199C-1086
5	11.8	59.0	17	6	US-09-817-879-1279
6	11.8	59.0	17	6	US-09-817-879-1504
7	11.8	59.0	17	6	US-09-817-879-1505
8	11.8	59.0	17	6	US-09-817-879-3051
9	11.8	59.0	17	6	US-09-817-879-3276
10	11.8	59.0	31	5	US-09-574-376B-157
11	11.8	59.0	31	7	US-09-724-671-4282
12	11.6	58.0	22	7	US-09-918-889-26
13	11.6	58.0	22	7	US-09-919-042-26
14	11.6	58.0	24	5	US-09-929-404-282
15	11.6	58.0	41	1	PCT-US01-25861-1
16	11.6	58.0	41	5	US-09-932-581-1
17	11.4	57.0	42	6	US-09-477-962-23
18	11.4	57.0	48	8	US-60-252-833-6500
19	11.2	56.0	20	6	US-09-800-629A-185
20	11.2	56.0	43	8	US-60-253-653-10876
21	11.2	56.0	46	8	US-60-253-378-27275
22	11	55.0	20	6	US-09-198-452A-6446
23	11	55.0	26	7	US-09-145-916B-22
24	11	55.0	31	5	US-09-574-376B-1105
25	11	55.0	31	6	US-09-817-879-6300

c 26	11	55.0	35	7	US-09-864-785-3865	Sequence 3865, Ap
c 27	11	55.0	37	6	US-09-595-639-6	Sequence 6, Appli
c 28	11	55.0	38	5	US-09-371-772B-12058	Sequence 12058, A
c 29	11	55.0	38	5	US-09-708-690-14335	Sequence 14335, A
c 30	11	55.0	41	8	US-60-253-653-23884	Sequence 23884, A
c 31	11	55.0	43	7	US-09-880-505-81	Sequence 81, Appl
c 32	11	55.0	45	7	US-09-633-739-10	Sequence 10, Appl
c 33	11	55.0	45	8	US-60-253-654-19893	Sequence 19893, A
c 34	11	55.0	45	8	US-60-255-592-19893	Sequence 19893, A
c 35	10.8	54.0	15	5	US-09-274-553C-355	Sequence 355, App
c 36	10.8	54.0	15	5	US-09-274-553C-1035	Sequence 1025, Ap
c 37	10.8	54.0	16	5	US-09-765-400-33	Sequence 33, Appl
c 38	10.8	54.0	17	6	US-09-817-879-1280	Sequence 1280, Ap
c 39	10.8	54.0	17	6	US-09-817-879-3050	Sequence 3050, Ap
c 40	10.8	54.0	17	6	US-09-817-879-3277	Sequence 3277, Ap
c 41	10.8	54.0	18	6	US-09-863-806-180	Sequence 180, App
c 42	10.8	54.0	18	6	US-09-863-806-186	Sequence 186, App
c 43	10.8	54.0	23	6	US-09-803-110-1474	Sequence 1474, Ap
c 44	10.8	54.0	32	1	PCT-US01-1337A-40	Sequence 40, Appl
c 45	10.8	54.0	36	7	US-09-846-040-165	Sequence 165, App

ALIGNMENTS

RESULT 1
US-60-252-833-26301
; Sequence 26301, Application US/60252833
; GENERAL INFORMATION:
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R
; TITLE OF INVENTION: Compositions isolated from bovine
; TITLE OF INVENTION: tissues and methods for their use.
; FILE REFERENCE: 1052P2
; CURRENT APPLICATION NUMBER: US/60/252,833
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 43535
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 26301
; LENGTH: 43
; TYPE: DNA
; ORGANISM: Bovine
US-60-252-833-26301

Query Match 61.0%; Score 12.2; DB 8; Length 43;
Best Local Similarity 82.4%; Pred. No. 9.1e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggcgcgtgaagcgaag 17
| | | | | | | | | | | | | | | | | | |
Db 1 ggcgcgtgaagcgaag 17

RESULT 2
US-09-475-947A-204
; Sequence 204, Application US/09475947A
; GENERAL INFORMATION:
; APPLICANT: Garner, Harold R.
; APPLICANT: Wren, Jonathan D.
; APPLICANT: Minna, John D.
; TITLE OF INVENTION: Polymorphic Repeats in Human Genes
; FILE REFERENCE: UTS00667
; CURRENT APPLICATION NUMBER: US/09/475,947A
; CURRENT FILING DATE: 1999-12-31
; NUMBER OF SEQ ID NOS: 346
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 204
; LENGTH: 35
; TYPE: DNA
; ORGANISM: human
US-09-475-947A-204

```
Query Match      60.0%; Score 12; DB 7; Length 35;
Best Local Similarity 75.0%; Pred. No. 1.1e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcggaagca 20
    ||| |||| ||| ||| |||
Db 8 ggcgcgtgaagcggaagca 27
    ||| |||| ||| ||| |||

RESULT 3
US-09-442-849B-1086
; Sequence 1086, Application US/09442849B
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Nucleic Acids Containing Single Nucleotide
; TITLE OF INVENTION: Polymorphisms and Methods of Use Thereof
; FILE REFERENCE: 15966-534C
; CURRENT APPLICATION NUMBER: US/09/442,849B
; PRIOR FILING DATE: 1999-11-17
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 09/443,199
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/442, 129
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 1592
; SOFTWARE: Curagen Patent Formatter Version 0.9
; SEQ ID NO 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (26)...(0)
; OTHER INFORMATION: 2 of 2 allelic variants (1085 is other entry)
; NAME/KEY: misc_feature
; LOCATION: (25)...(26)
; OTHER INFORMATION: nucleotide deleted between bases 25 and 26
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Accession number cg44923068
; US-09-442-849B-1086

Query Match      60.0%; Score 12; DB 7; Length 50;
Best Local Similarity 75.0%; Pred. No. 1.2e+04;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcggaagca 20
    ||| |||| ||| ||| |||
Db 23 ggactcgtgaagcggaagca 42
    ||| |||| ||| ||| |||

RESULT 5
US-09-817-879-1279/c
; Sequence 1279, Application US/09817879
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Rel
; FILE REFERENCE: MBH00-801-F
; CURRENT APPLICATION NUMBER: US/09/817,879
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9703
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1279
; LENGTH: 17
; TYPE: RNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: oligonucleotide substrate
; US-09-817-879-1279

Query Match      59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 86.7%; Pred. No. 1.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ccgtgaagcggaagcc 19
    |||| |||| ||||
Db 17 CCGGTAGCGTAGGC 3
    |||| |||| ||||

RESULT 6
US-09-817-879-1504/c
; Sequence 1504, Application US/09817879
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals Inc.
; TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Rel
; FILE REFERENCE: MBH00-801-F
; CURRENT APPLICATION NUMBER: US/09/817,879
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 9703
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1504
; LENGTH: 17
; TYPE: RNA
; ORGANISM: artificial sequence
; FEATURE:
; NAME/KEY: misc_feature
```

LOCATION:
OTHER INFORMATION: oligonucleotide substrate
US-09-817-879-1504

Query Match 59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 86.7%; Pred. No. 1.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ccgtgaagcggaagc 19
||| ||||| |||
Db 17 CCGGAAGCGAAGC 3

RESULT 7

US-09-817-879-1505/c
Sequence 1505, Application US/09817879
GENERAL INFORMATION:
APPLICANT: Ribozyme Pharmaceuticals Inc.
TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to Hepatitis C Virus Infection
FILE REFERENCE: MH800-801-F
CURRENT APPLICATION NUMBER: US/09/817,879
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 9703
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1505
LENGTH: 17
TYPE: RNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: oligonucleotide substrate
US-09-817-879-1505

Query Match 59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 86.7%; Pred. No. 1.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 gcgcgtgaagcgaa 16
||| ||||| |||
Db 15 GCCCGGAGCGGAA 1

RESULT 8

US-09-817-879-3051
Sequence 3051, Application US/09817879
GENERAL INFORMATION:
APPLICANT: Ribozyme Pharmaceuticals Inc.
TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to Hepatitis C Virus Infection
FILE REFERENCE: MH800-801-F
CURRENT APPLICATION NUMBER: US/09/817,879
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 9703
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3051
LENGTH: 17
TYPE: RNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: oligonucleotide substrate
US-09-817-879-3051

Query Match 59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 86.7%; Pred. No. 1.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ccgtgaagcggaagc 19
||| ||||| |||
Db 2 ccggaagcggaagc 16

RESULT 9

US-09-817-879-3276
Sequence 3276, Application US/09817879
GENERAL INFORMATION:
APPLICANT: Ribozyme Pharmaceuticals Inc.
TITLE OF INVENTION: Enzymatic Nucleic Acid Treatment of Diseases or Conditions Related to Hepatitis C Virus Infection
FILE REFERENCE: MH800-801-F
CURRENT APPLICATION NUMBER: US/09/817,879
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 9703
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3276
LENGTH: 17
TYPE: RNA
ORGANISM: artificial sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION:
OTHER INFORMATION: oligonucleotide substrate
US-09-817-879-3276

Query Match 59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 80.0%; Pred. No. 1.4e+04;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ccgtgaagcggaagc 19
||| ||||| |||
Db 2 ccggaagcggaagc 16

RESULT 10

US-09-574-376B-157
Sequence 157, Application US/09574376B
GENERAL INFORMATION:
APPLICANT: Warrington, Janet
APPLICANT: Shah, Nila
APPLICANT: Gingeras, Thomas Raymond
APPLICANT: Affymetrix, Inc.
TITLE OF INVENTION: Human Lymphoblast Polymorphisms
FILE REFERENCE: 3229.2
CURRENT APPLICATION NUMBER: US/09/574,376B
CURRENT FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 1330
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 157
LENGTH: 31
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
US-09-574-376B-157

Query Match 59.0%; Score 11.8; DB 5; Length 31;
Best Local Similarity 86.7%; Pred. No. 1.4e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 cgtgaagcggaagca 20
||| ||||| |||||
Db 17 cgagaagagaagca 31

RESULT 11

US-09-724-671-4282/c
Sequence 4282, Application US/09724671
GENERAL INFORMATION:
APPLICANT: Watson, James D

```

; APPLICANT: Murison, James G
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed
; TITLE OF INVENTION: by the polynucleotides and methods for their use.
; FILE REFERENCE: 105002
; CURRENT APPLICATION NUMBER: US/09/724.671
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: U.S. No. 60/171,678
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 21907
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4282
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Mouse
; US-09-724-671-4282

```

```

RESULT 12
US-09-918-889-26/c
; Sequence 26, Application US/09918889
; GENERAL INFORMATION:
; APPLICANT: Readhead, Carol W.
; APPLICANT: Winston, Robert
; APPLICANT: Koefler, H. Phillip
; APPLICANT: Muller, Carsten
; TITLE OF INVENTION: NUCLEIC ACID CONSTRUCTS CONTAINING A
; TITLE OF INVENTION: CYCLIN AL PROMOTER, AND KIT
; FILE REFERENCE: 18810-81603
; CURRENT APPLICATION NUMBER: US/09/918,889
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US 09/292,723
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: US 09/191,920
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 60/065,825
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/US98/24238
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Single-stranded oligonucleotide
US-09-918-889-26

```

```

Query Match      58.0%; Score 11.6; DB 7; Length 22;
Best Local Similarity 77.8%; Pred. No. 1.7e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1  ggcgccatgaagcgaag 18
         ||||| ||||| |||||
Db      22  GGGCGAGGAAGGGAAGG 5

RESULT 13
US-09-919-042-26/c
; Sequence 26, Application US/09919042
; GENERAL INFORMATION:
; APPLICANT: Readhead, Carol W.
; APPLICANT: Winston, Robert

```

```

; APPLICANT: Koeffler, H. Phillip
; APPLICANT: Muller, Carsten
; TITLE OF INVENTION: Transfection, Storage and Transfer of
; TITLE OF INVENTION: Male Germ Cells for Generation of Selectable Transgenic Stem
; TITLE OF INVENTION: Cells
; FILE REFERENCE: 18810-81602
; CURRENT APPLICATION NUMBER: US/09/919,042
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US 09/292,723
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: US 09/191,920
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: US 60/065,825
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: PCT/US98/24238
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Single-stranded oligonucleotide
; 95-09-919-042-26

```

```

RESULT 14
US-09-929-404-282
; Sequence 282, Application US/09929404
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hebert, Caroline
; APPLICANT: Hensel, William
; APPLICANT: Kabokoff, Rhona C.
; APPLICANT: Lu, Yanmei
; APPLICANT: Pan, James
; APPLICANT: Pennica, Diane
; APPLICANT: Shelton, David L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William
; APPLICANT: Yan, Minhong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE
; TITLE OF INVENTION: RELATED DISEASES
; FILE REFERENCE: P2833R1C1
; CURRENT APPLICATION NUMBER: US/09/929,404
; CURRENT FILING DATE: 2001-08-13
; NUMBER OF SEQ ID NOS: 310
; SEQ ID NO 282
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: forward PCR primer
US-09-929-404-282
Query Match 58.0%; Score 11.6; DB 5; Length 24;

```

Best Local Similarity 77.8%; Pred. No. 1.8e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggcgcgtgaagcgaagg 18
| ||| ||||| |||
Db 6 gtcgctgtgaagccaacg 23

RESULT 15
PCT-US01-25861-1/c
; Sequence 1, Application PC/TUS0125861
; GENERAL INFORMATION:
; APPLICANT: Sierra Sciences, Inc.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR MODULATING
; FILE REFERENCE: SIER-005WO
; CURRENT APPLICATION NUMBER: PCT/US01/25861
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/227,865
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/230,174
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/238,345
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 41
; TYPE: DNA
; ORGANISM: human
PCT-US01-25861-1

Query Match 58.0%; Score 11.6; DB 1; Length 41;
Best Local Similarity 77.8%; Pred. No. 1.8e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ggcgcgtgaagcgaaggc 19
| |||| | | | | |
Db 24 GCCTCGGAGGAGAGGCG 7

Search completed: October 2, 2001, 16:54:59
Job time: 17662 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 21:50:08 ; Search time 17695.9 Seconds
(without alignments)
16.701 Million cell updates/sec

Title: US-09-757-100B-4

Perfect score: 20

Sequence: 1 ggcgcgctgaagcgaaggca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 14155048 seqs, 7388405095 residues

Total number of hits satisfying chosen parameters: 4519004

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US098_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US099_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US099H_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US099I_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US099J_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US099K_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	20	100.0	20	1	PCT-US00-18999-4
2	20	100.0	20	17	US-09-377-310-4
3	20	100.0	20	29	US-09-757-100B-4
4	15	75.0	15	1	PCT-US00-18999-24
5	15	75.0	15	17	US-09-377-310-4
6	15	75.0	15	29	US-09-757-100B-24
7	14.8	74.0	45	8	US-08-448-223-13
8	14.8	74.0	45	8	US-08-448-223-14
9	14.4	72.0	25	55	US-60-234-017-94926
10	13.6	68.0	25	55	US-60-234-017-952075
11	13.6	68.0	30	31	US-09-812-913-11
12	13.6	68.0	41	18	US-09-404-520-40172
13	13.6	68.0	41	18	US-09-404-520-40466
14	13.2	66.0	25	55	US-60-232-638-45792
15	13.2	66.0	25	55	US-60-233-166-104969
16	13.2	66.0	25	55	US-60-233-166-104976
17	13.2	66.0	25	55	US-60-233-166-104980
18	13.2	66.0	25	55	US-60-233-166-105010
19	13	65.0	25	55	US-60-234-017-568725
20	13	65.0	25	55	US-60-234-017-568728
21	13	65.0	25	55	US-60-234-017-568731
22	13	65.0	25	55	US-60-234-017-568734
23	12.8	64.0	25	26	US-09-660-080-19599
24	12.8	64.0	25	26	US-09-660-220-12063
25	12.8	64.0	25	55	US-60-233-166-200509
26	12.8	64.0	25	55	US-60-233-166-309711
27	12.8	64.0	25	55	US-60-233-620-2593
28	12.8	64.0	25	55	US-60-233-620-78332
29	12.8	64.0	25	55	US-60-233-620-83352
30	12.8	64.0	25	55	US-60-233-620-83352
31	12.8	64.0	25	55	US-60-234-049-4741
32	12.8	64.0	25	55	US-60-234-049-4743
33	12.8	64.0	25	55	US-60-234-049-4744
34	12.8	64.0	25	55	US-60-234-049-4746
35	12.8	64.0	25	55	US-60-234-049-4754
36	12.8	64.0	25	55	US-60-234-049-4788
37	12.8	64.0	25	55	US-60-234-049-4789
38	12.8	64.0	25	55	US-60-234-049-4790
39	12.8	64.0	25	55	US-60-234-049-4791
40	12.8	64.0	25	55	US-60-234-049-4792
41	12.8	64.0	25	55	US-60-234-049-4793
42	12.8	64.0	25	55	US-60-234-049-4795
43	12.8	64.0	25	55	US-60-234-049-5481
44	12.8	64.0	25	55	US-60-234-049-59410
45	12.8	64.0	25	55	US-60-234-049-73099

ALIGNMENTS

```
RESULT 1
PCT-US00-18999-4
; Sequence 4, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-4

Query Match      100.0%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcggaagca 20
Db 1 ggcgcgtgaagcggaagca 20
|||||

RESULT 2
US-09-377-310-4
; Sequence 4, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-4

Query Match      100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcggaagca 20
Db 1 ggcgcgtgaagcggaagca 20
|||||

RESULT 3
US-09-757-100B-4
; Sequence 4, Application US/09757100B
```

```
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-4

Query Match      100.0%; Score 20; DB 29; Length 20;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcggaagca 20
Db 1 ggcgcgtgaagcggaagca 20
|||||

RESULT 4
PCT-US00-18999-24
; Sequence 24, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-24

Query Match      75.0%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 cgcctgaagcggaag 17
Db 1 cgcctgaagcggaag 15
|||||

RESULT 5
US-09-377-310-24
; Sequence 24, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
```



```

; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-24

```

```

Query Match          75.0%; Score 15; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 cgcgcgtgaagcgaag 17
    |||||
DB 1 cgcgcgtgaagcgaag 15

```

```

RESULT 6
US-09-757-100B-24
; Sequence 24, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-24

```

```

Query Match          75.0%; Score 15; DB 29; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 cgcgcgtgaagcgaag 17
    |||||
DB 1 cgcgcgtgaagcgaag 15

```

```

RESULT 7
US-08-448-223-13
; Sequence 13, Application US/08448223
; GENERAL INFORMATION:
; APPLICANT: Abramson, Richard D.
; APPLICANT: Gelfand, David H.
; APPLICANT: Kalman, Lisa V.
; APPLICANT: Reichert, Fred L.
; TITLE OF INVENTION: Thermostable DNA Polymerases With
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/08/448,223
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-08-448-223-13

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Sias, Stacey R.
; REGISTRATION NUMBER: 32,630
; REFERENCE/DOCKET NUMBER: 9134
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 814-2863
; TELEFAX: (510) 814-2977
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-448-223-13

```

```

Query Match          74.0%; Score 14.8; DB 8; Length 45;
Best Local Similarity 88.9%; Pred. No. 4.9e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 2 gcgcgcgtgaagcgaagc 19
    |||||
DB 3 gcgcgcgtgaagcgaagc 20

```

```

RESULT 8
US-08-448-223-14
; Sequence 14, Application US/08448223
; GENERAL INFORMATION:
; APPLICANT: Abramson, Richard D.
; APPLICANT: Gelfand, David H.
; APPLICANT: Kalman, Lisa V.
; APPLICANT: Reichert, Fred L.
; TITLE OF INVENTION: Thermostable DNA Polymerases With
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/08/448,223
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-08-448-223-14

```

```

Query Match          75.0%; Score 15; DB 29; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 cgcgcgtgaagcgaag 17
    |||||
DB 1 cgcgcgtgaagcgaag 15

```

```

; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; : DATABASE ACCESSION NUMBER: GenBank AI508367
US-60-234-017-352075

```

```

Query Match      68.0%; Score 13.6; DB 18; Length 41;
Best Local Similarity 80.0%; Pred. No. 1.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1  ggccgcgtgaagcgaaggca 20
          ||||| ||| ||| ||| |||
Db      13  gccaccatgaqgcaaaaggca 32

```

```
RESULT 13
US-09-404-520-40466
; Sequence 40466, Application US/09404520
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Ghodssi, Azita
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: McIninch, James
; APPLICANT: Timberlake, William E.
; APPLICANT: Yu, Jaehyuk
; TITLE OF INVENTION: Emmericella nidulans Genome Sequence and Uses Thereof
; FILE REFERENCE: 38-10(15498)A
; CURRENT APPLICATION NUMBER: US/09/404,520
; CURRENT FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 44345
; SEQ ID NO 40466
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Aspergillus nidulans
US-09-404-520-40466

Query Match      68.0%; Score 13.6; DB 18; Length 41;
Best Local Similarity 80.0%; Pred. No. 1.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ggcgcgcgtgaagcgaaaggca 20
   | |||| |||| | ||||
Db 13 ggcgcctgaagaggaggagca 32

RESULT 14
US-60-232-638-45792
; Sequence 45792, Application US/60232638
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; TITLE OF INVENTION: Methods of Genetic Analysis of Yeast
; FILE REFERENCE: 3110
; CURRENT APPLICATION NUMBER: US/60/232,638
; CURRENT FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 138410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45792
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: SGD YDL200C
US-60-232-638-45792

Query Match      66.0%; Score 13.2; DB 55; Length 25;
Best Local Similarity 83.3%; Pred. No. 2.9e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 cgcgcgtgaagcgaaaggca 20
   | |||| | |||| |
Db 5 cgcgcgtcaaccgaaaggga 22

RESULT 15
US-60-233-166-104969/c
; Sequence 104969, Application US/60233166
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat
; FILE REFERENCE: 3112
; CURRENT APPLICATION NUMBER: US/60/233,166
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 104969
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AA859343
US-60-233-166-104969
```

```
Query Match      66.0%; Score 13.2; DB 55; Length 25;
Best Local Similarity 83.3%; Pred. No. 2.9e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ggcgcgtgaagcgaaaggc 19
   | |||| | |||| |
Db 24 GGGCCGCTAGCGAAGGC 7
```

```
Search completed: October 2, 2001, 21:50:08
Job time: 24521 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 15:01:27 ; Search time 17695.9 Seconds
(without alignments)
16.701 Million cell updates/sec

Title: US-09-757-100B-3
Perfect score: 20
Sequence: 1 ccgcgggtcacagtgtcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1415048 seqs, 7388405095 residues
Total number of hits satisfying chosen parameters: 4519004

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/US086 COMB.seq:*

2: /cgn2_6/ptodata/1/pna/US07 COMB.seq:*

3: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US08 COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	20	100.0	20	1	PCT-US00-18999-3	Sequence 3, Appli
2	20	100.0	20	17	US-09-377-310-3	Sequence 3, Appli
3	20	100.0	20	17	US-09-377-310-3	Sequence 3, Appli
4	15	75.0	15	1	PCT-US00-18999-23	Sequence 23, Appl
5	15	75.0	15	17	US-09-377-310-23	Sequence 23, Appl
6	15	75.0	15	29	US-09-757-100B-23	Sequence 23, Appl
7	14.8	74.0	25	26	US-09-660-220-37376	Sequence 37376, A
8	14.8	74.0	25	26	US-09-660-220-37377	Sequence 37377, A
9	14.8	74.0	25	55	US-60-233-166-185790	Sequence 185790, A
10	14.2	71.0	25	55	US-60-234-017-55657	Sequence 55657, A
11	14.2	71.0	25	55	US-60-234-017-75527	Sequence 75527, A
12	14	70.0	25	55	US-60-233-166-279728	Sequence 279728, A
13	13.8	69.0	25	55	US-60-233-166-319513	Sequence 319513, A
14	13.8	69.0	25	55	US-60-234-017-10846	Sequence 10846, A
15	13.8	69.0	25	55	US-60-234-017-248844	Sequence 248844, A
16	13.6	68.0	40	9	US-08-594-866-2	Sequence 2, Appli
17	13.6	68.0	25	55	US-60-233-166-418139	Sequence 418139, A
18	13.6	68.0	35	29	US-09-724-866-18962	Sequence 18962, A
19	13.6	68.0	35	29	US-09-724-866A-18962	Sequence 18962, A
20	13.6	68.0	35	49	US-60-171-432-18962	Sequence 18962, A
21	13.6	68.0	39	14	US-09-083-194A-6	Sequence 6, Appli
22	13.6	68.0	41	17	US-09-326-203A-31	Sequence 31, Appli
23	13.4	67.0	25	55	US-60-232-638-136549	Sequence 136549, A
24	13.4	67.0	25	55	US-60-233-166-296555	Sequence 296555, A
25	13.4	67.0	25	55	US-60-234-017-217856	Sequence 217856, A
26	13.2	66.0	25	55	US-60-233-166-410983	Sequence 410983, A
27	13.2	66.0	25	55	US-60-233-166-413618	Sequence 413618, A
28	13.2	66.0	25	55	US-60-234-017-125344	Sequence 125344, A
29	13.2	66.0	29	17	US-09-304-232-546	Sequence 242830, A
30	13.2	66.0	31	14	US-09-006-298-11	Sequence 546, Appl
31	13.2	66.0	39	18	US-09-405-459-173	Sequence 173, Appl
32	13.2	66.0	50	21	US-09-549-848A-57	Sequence 57, Appl
33	13.2	66.0	50	21	US-09-549-848B-57	Sequence 57, Appl
34	12.8	64.0	25	26	US-09-660-220-37375	Sequence 37375, A
35	12.8	64.0	25	55	US-60-233-166-30520	Sequence 30520, A
36	12.8	64.0	25	55	US-60-233-166-40877	Sequence 40877, A
37	12.8	64.0	25	55	US-60-233-166-185404	Sequence 185404, A
38	12.8	64.0	25	55	US-60-233-166-235736	Sequence 235736, A
39	12.8	64.0	25	55	US-60-233-166-272127	Sequence 272127, A
40	12.8	64.0	25	55	US-60-233-166-272785	Sequence 272785, A
41	12.8	64.0	25	55	US-60-233-166-341785	Sequence 341785, A
42	12.8	64.0	25	55	US-60-233-166-341789	Sequence 341789, A
43	12.8	64.0	25	55	US-60-233-166-404414	Sequence 404414, A
44	12.8	64.0	25	55	US-60-234-017-480869	Sequence 480869, A

ALIGNMENTS

RESULT 1
PCT-US00-18999-3
; Sequence 3, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-3

Query Match 100.0%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccgcgggctcacagtgtcg 20
Db 1 ccgcgggctcacagtgtcg 20

RESULT 2
US-09-377-310-3
; Sequence 3, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-3

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccgcgggctcacagtgtcg 20
Db 1 ccgcgggctcacagtgtcg 20

RESULT 3
US-09-757-100B-3
; Sequence 3, Application US/09757100B

; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-3

Query Match 100.0%; Score 20; DB 29; Length 20;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccgcgggctcacagtgtcg 20
Db 1 ccgcgggctcacagtgtcg 20

RESULT 4
PCT-US00-18999-23
; Sequence 23, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-23

Query Match 75.0%; Score 15; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 gcgggctcacagtgg 17
Db 1 gcgggctcacagtgg 15

RESULT 5
US-09-377-310-23
; Sequence 23, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.

; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-23

Query Match 75.0%; Score 15; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gcggggctcacagtgg 17
| | | | | | | | | | | | | | |
Db 1 gcggggctcacagtgg 15

RESULT 6
US-09-757-100B-23
; Sequence 23, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-23

Query Match 75.0%; Score 15; DB 29; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.1e+03;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 gcggggctcacagtgg 17
| | | | | | | | | | | | | | |
Db 1 gcggggctcacagtgg 15

RESULT 7
US-09-660-220-37376
; Sequence 37376, Application US/09660220
; GENERAL INFORMATION:
; APPLICANT: Mittmann et al.
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of Human
; FILE REFERENCE: 3102.1
; CURRENT APPLICATION NUMBER: US/09/660,220
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/164,973

; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 140981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37376
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo Sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank L11931
US-09-660-220-37376

Query Match 74.0%; Score 14.8; DB 26; Length 25;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cgcgggctcacagtggtc 19
| | | | | | | | | | | | | | |
Db 6 cccgggctcacagtgtgc 23

RESULT 8
US-09-660-220-37377
; Sequence 37377, Application US/09660220
; GENERAL INFORMATION:
; APPLICANT: Mittmann et al.
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of Human
; FILE REFERENCE: 3102.1
; CURRENT APPLICATION NUMBER: US/09/660,220
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: 60/164,973
; PRIOR FILING DATE: 1999-11-11
; NUMBER OF SEQ ID NOS: 140981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37377
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo Sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank L11931
US-09-660-220-37377

Query Match 74.0%; Score 14.8; DB 26; Length 25;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cgcgggctcacagtggtc 19
| | | | | | | | | | | | | | |
Db 2 cccgggctcacagtgtgc 19

RESULT 9
US-60-233-166-185790
; Sequence 185790, Application US/60233166
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat
; FILE REFERENCE: 3112
; CURRENT APPLICATION NUMBER: US/60/233,166
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 185790
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AA963963
US-60-233-166-185790

Query Match 74.0%; Score 14.8; DB 55; Length 25;
Best Local Similarity 88.9%; Pred. No. 5.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 gcggggtcacagtgtgcg 20
||||| ||||| ||||| |||||
Db 2 gcggactcacgtgtgcg 19

RESULT 10
US-60-234-017-55657
; Sequence 55657, Application US/60234017
; GENERAL INFORMATION:
; APPLICANT: Mittmann, M
; TITLE OF INVENTION: Methods of Genetic Analysis of Mus
; FILE REFERENCE: 3115
; CURRENT APPLICATION NUMBER: US/60/234,017
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 605887
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 55657
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank U76759
US-60-234-017-55657

Query Match 71.0%; Score 14.2; DB 55; Length 25;
Best Local Similarity 84.2%; Pred. No. 1e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 cgcgggtcacagtgtgcg 20
|| ||||| ||||| |||||
Db 2 cgtgggtcacagagtgcg 20

RESULT 11
US-60-234-017-75527/c
; Sequence 75527, Application US/60234017
; GENERAL INFORMATION:
; APPLICANT: Mittmann, M
; TITLE OF INVENTION: Methods of Genetic Analysis of Mus
; FILE REFERENCE: 3115
; CURRENT APPLICATION NUMBER: US/60/234,017
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 605887
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 75527
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AF048838
US-60-234-017-75527

Query Match 71.0%; Score 14.2; DB 55; Length 25;
Best Local Similarity 84.2%; Pred. No. 1e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccgcgggtcacagtgtgcg 19
||||| ||||| ||||| |||||
Db 19 CCGCGGACCCACAGTGGAC 1

RESULT 12
US-60-233-166-279728/c
; Sequence 279728, Application US/60233166
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat
; FILE REFERENCE: 3112
; CURRENT APPLICATION NUMBER: US/60/233,166
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 279728
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AF639107
US-60-233-166-279728

Query Match 70.0%; Score 14; DB 55; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ggtcacagtgtgc 19
||||| ||||| ||||| |||||
Db 21 GGTCACACAGTGTGC 8

RESULT 13
US-60-233-166-319513
; Sequence 319513, Application US/60233166
; GENERAL INFORMATION:
; APPLICANT: Mittmann
; TITLE OF INVENTION: Methods of Genetic Analysis of the Rat
; FILE REFERENCE: 3112
; CURRENT APPLICATION NUMBER: US/60/233,166
; CURRENT FILING DATE: 2000-10-24
; NUMBER OF SEQ ID NOS: 420907
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 319513
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank L02896
US-60-233-166-319513

Query Match 69.0%; Score 13.8; DB 55; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 cggggtcacagtgtgcg 20
||||| ||||| ||||| |||||
Db 2 cggactcacagtgtgcg 18

RESULT 14
US-60-234-017-10846
; Sequence 10846, Application US/60234017
; GENERAL INFORMATION:
; APPLICANT: Mittmann, M
; TITLE OF INVENTION: Methods of Genetic Analysis of Mus
; FILE REFERENCE: 3115
; CURRENT APPLICATION NUMBER: US/60/234,017
; CURRENT FILING DATE: 2000-09-20

; NUMBER OF SEQ ID NOS: 605887
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10846
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AF055638
US-60-234-017-10846

Query Match 69.0%; Score 13.8; DB 55; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cgggggtcacagtgg 18
||| |||||
Db 6 cgggtctcacagtgg 22

RESULT 15
US-60-234-017-248844
; Sequence 248844, Application US/60234017
; GENERAL INFORMATION:
; APPLICANT: Mittmann, M
; TITLE OF INVENTION: Methods of Genetic Analysis of Mus
; TITLE OF INVENTION: musculus
; FILE REFERENCE: 3115
; CURRENT APPLICATION NUMBER: US/60/234,017
; CURRENT FILING DATE: 2000-09-20
; NUMBER OF SEQ ID NOS: 605887
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 248844
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank AW049900
US-60-234-017-248844

Query Match 69.0%; Score 13.8; DB 55; Length 25;
Best Local Similarity 88.2%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 cgggggtcacagtgg 20
||| |||||
Db 7 cgggggtcacagtgg 23

Search completed: October 2, 2001, 21:50:08
Job time: 24521 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 12:00:37 ; Search time 2173.58 Seconds
(without alignments)
19.290 Million cell updates/sec

Title: US-09-757-100b-3

Perfect score: 20

Sequence: 1 ccgcgggtcacagtgtcg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1895189 seqs, 1048201267 residues

Total number of hits satisfying chosen parameters: 299022

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq1.*
- 7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2.*
- 8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	13.2	66.0	50	7	US-09-688-069-57
2	13.2	66.0	50	7	US-09-688-071-57
3	12.2	61.0	41	7	US-09-688-069-51
4	12.2	61.0	41	7	US-09-688-071-51
5	12.2	61.0	45	7	US-09-688-069-53
6	12.2	61.0	45	7	US-09-688-071-53
7	11.8	59.0	25	7	US-09-735-271-2037
8	11.6	58.0	23	6	US-09-879-341-12
9	11.6	58.0	26	6	US-09-879-341-10
10	11.6	58.0	30	7	US-09-915-060-9
11	11.6	58.0	30	7	US-09-915-060-43
12	11.6	58.0	39	8	US-60-298-340-21
13	11.4	57.0	18	7	US-09-787-252-17
14	11.4	57.0	31	6	US-09-801-274-629
15	11.4	57.0	40	7	US-09-735-271-1558
16	11.4	57.0	45	6	US-09-336-643-50
17	11.2	56.0	17	5	US-09-371-772B-4196
18	11.2	56.0	17	5	US-09-708-690-4196
19	11.2	56.0	20	7	US-09-825-497-23
20	11.2	56.0	22	6	US-09-883-152-55
21	11.2	56.0	31	6	US-09-801-274-1447
22	11.2	56.0	38	7	US-09-688-069-55
23	11.2	56.0	38	7	US-09-688-069-59
24	11.2	56.0	38	7	US-09-688-071-55
25	11.2	56.0	38	7	US-09-688-071-59

c	26	11.2	56.0	40	5	US-09-941-992-324	Sequence 324, App
	27	11.2	56.0	47	1	PCT-US01-23545-17	Sequence 17, Appl
	28	11.2	56.0	47	1	PCT-US01-23707-17	Sequence 17, Appl
	29	11.2	56.0	47	7	US-09-912-436-17	Sequence 17, Appl
c	30	11.2	56.0	48	6	US-09-803-110-7471	Sequence 7471, Ap
c	31	11.2	56.0	48	7	US-09-724-750-7833	Sequence 7833, Ap
	32	11	55.0	20	6	US-09-851-896-76	Sequence 76, Appl
c	33	11	55.0	22	7	US-09-748-537-5	Sequence 5, Appli
c	34	11	55.0	37	6	US-09-803-110-1481	Sequence 1481, Ap
c	35	11	55.0	47	8	US-60-253-457-20631	Sequence 20631, A
c	36	11	55.0	49	8	US-60-252-833-32799	Sequence 32799, A
	37	11	55.0	50	6	US-09-504-576A-10254	Sequence 10254, A
c	38	10.8	54.0	16	5	US-09-371-772B-5913	Sequence 5913, Ap
c	39	10.8	54.0	16	5	US-09-708-690-5913	Sequence 5913, Ap
c	40	10.8	54.0	17	5	US-09-371-772B-5002	Sequence 5002, Ap
c	41	10.8	54.0	17	5	US-09-708-690-5002	Sequence 5002, Ap
c	42	10.8	54.0	17	5	US-09-708-690-7832	Sequence 7832, Ap
c	43	10.8	54.0	17	6	US-09-546-745A-2312	Sequence 2312, Ap
c	44	10.8	54.0	17	6	US-09-817-879-143	Sequence 143, App
c	45	10.8	54.0	17	6	US-09-817-879-4412	Sequence 4412, Ap

ALIGNMENTS

RESULT 1
US-09-688-069-57
; Sequence 57, Application US/09688069
; GENERAL INFORMATION:
; APPLICANT: Subramaniam, S.; Slater, S.; Karberg, K.; Chen, R.; Valentin, H.; Wong,
; TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synt
; FILE REFERENCE: 16515.054
; CURRENT APPLICATION NUMBER: US/09/688,069
; CURRENT FILING DATE: 2000-10-14
; NUMBER OF SEQ ID NOS: 114
; SEQ ID NO 57
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotides
US-09-688-069-57

Query Match 66.0% Score 13.2; DB 7; Length 50;
Best Local Similarity 83.3%; Pred. No. 2.9e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccgcgggtcacagtgtg 18
Db 5 ccgcgggtcacagtgtg 22

RESULT 2
US-09-688-071-57
; Sequence 57, Application US/09688071
; GENERAL INFORMATION:
; APPLICANT: Lassner, M.; Post-Beittenmiller, M.; Savidge, B.; Weiss, J.; Mitsky, T.
; TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synt
; FILE REFERENCE: 16515.055
; CURRENT APPLICATION NUMBER: US/09/688,071
; CURRENT FILING DATE: 2000-10-14
; PRIOR APPLICATION NUMBER: US 60/129,899
; PRIOR FILING DATE: 1999-04-15
; PRIOR APPLICATION NUMBER: US 60/146,461
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: US 09/549,848
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: WO PCT/US00/10368
; PRIOR FILING DATE: 2000-09-14
; NUMBER OF SEQ ID NOS: 114
; SEQ ID NO 57

; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotides
US-09-688-071-57

Query Match 66.0%; Score 13.2; DB 7; Length 50;

Best Local Similarity 83.3%; Pred. No. 2.9e+03;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccgcggggtcacagtgtt 18

||||| | |||| | ||

Db 5 ccgcggccgcacaatggt 22

RESULT 3

US-09-688-069-51

; Sequence 51, Application US/09688069

; GENERAL INFORMATION:

; APPLICANT: Subramaniam, S.; Slater, S.; Karberg, K.; Chen, R.; Valentin, H.; Wong, Y.

; TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synthes

; FILE REFERENCE: 16515.054

; CURRENT APPLICATION NUMBER: US/09/688.069

; CURRENT FILING DATE: 2000-10-14

; NUMBER OF SEQ ID NOS: 114

; SEQ ID NO 51

; LENGTH: 41

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic Oligonucleotides

US-09-688-069-51

Query Match 61.0%; Score 12.2; DB 7; Length 41;

Best Local Similarity 82.4%; Pred. No. 9.4e+03;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccgcggggtcacagtgg 17

||||| | |||| | ||

Db 5 ccgcggccgcacaatgg 21

RESULT 4

US-09-688-071-51

; Sequence 51, Application US/09688071

; GENERAL INFORMATION:

; APPLICANT: Lassner, M.; Post-Beittenmiller, M.; Savidge, B.; Weiss, J.; Mitsky, T.;

; APPLICANT: Valentin, H.

; TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synthes

; FILE REFERENCE: 16515.055

; CURRENT APPLICATION NUMBER: US/09/688.071

; CURRENT FILING DATE: 2000-10-14

; PRIOR APPLICATION NUMBER: US 60/129,899

; PRIOR FILING DATE: 1999-04-15

; PRIOR APPLICATION NUMBER: US 60/146,461

; PRIOR FILING DATE: 1999-07-30

; PRIOR APPLICATION NUMBER: US 09/549,848

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: WO PCT/US00/10368

; PRIOR FILING DATE: 2000-09-14

; NUMBER OF SEQ ID NOS: 114

; SEQ ID NO 51

; LENGTH: 41

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic Oligonucleotides

US-09-688-071-51

Query Match 61.0%; Score 12.2; DB 7; Length 41;

Best Local Similarity 82.4%; Pred. No. 9.4e+03;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccgcggggtcacagtgg 17

||||| | |||| | ||

Db 5 ccgcggccgcacaatgg 21

RESULT 5

US-09-688-069-53

; Sequence 53, Application US/09688069

; GENERAL INFORMATION:

; APPLICANT: Subramaniam, S.; Slater, S.; Karberg, K.; Chen, R.; Valentin, H.; Wong,

; TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synt

; FILE REFERENCE: 16515.054

; CURRENT APPLICATION NUMBER: US/09/688.069

; CURRENT FILING DATE: 2000-10-14

; NUMBER OF SEQ ID NOS: 114

; SEQ ID NO 53

; LENGTH: 45

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic Oligonucleotides

US-09-688-069-53

Query Match 61.0%; Score 12.2; DB 7; Length 45;

Best Local Similarity 82.4%; Pred. No. 9.4e+03;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccgcggggtcacagtgg 17

||||| | |||| | ||

Db 5 ccgcggccgcacaatgg 21

RESULT 6

US-09-688-071-53

; Sequence 53, Application US/09688071

; GENERAL INFORMATION:

; APPLICANT: Lassner, M.; Post-Beittenmiller, M.; Savidge, B.; Weiss, J.; Mitsky, T.

; APPLICANT: Valentin, H.

; TITLE OF INVENTION: Nucleic Acid Sequences to Proteins Involved in Tocopherol Synt

; FILE REFERENCE: 16515.055

; CURRENT APPLICATION NUMBER: US/09/688.071

; CURRENT FILING DATE: 2000-10-14

; PRIOR APPLICATION NUMBER: US 60/129,899

; PRIOR FILING DATE: 1999-04-15

; PRIOR APPLICATION NUMBER: US 60/146,461

; PRIOR FILING DATE: 1999-07-30

; PRIOR APPLICATION NUMBER: US 09/549,848

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: WO PCT/US00/10368

; PRIOR FILING DATE: 2000-09-14

; NUMBER OF SEQ ID NOS: 114

; SEQ ID NO 53

; LENGTH: 45

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic Oligonucleotides

US-09-688-071-53

Query Match 61.0%; Score 12.2; DB 7; Length 45;

Best Local Similarity 82.4%; Pred. No. 9.4e+03;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccgcggggtcacagtgg 17

||||| | |||| | ||

Db 5 ccgcggccgcacaatgg 21

RESULT 7

US-09-735-271-2037
; Sequence 2037, Application US/09735271
; GENERAL INFORMATION:
; APPLICANT: Daly, Mark J.
; APPLICANT: Hudson, Thomas J.
; APPLICANT: Lander, Eric S.
; APPLICANT: Rioux, John
; APPLICANT: Siminovich, Kathy
; TITLE OF INVENTION: ISD-RELATED POLYMORPHISMS
; FILE REFERENCE: 2825.1025-002
; CURRENT APPLICATION NUMBER: US/09/735,271
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/170,257
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 60/196,046
; PRIOR FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 2058
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2037
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-271-2037

Query Match 59.0%; Score 11.8; DB 7; Length 25;

Best Local Similarity 86.7%; Pred. No. 1.5e+04; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 2;

Qy 2 cgcgggtcacagtg 16
| | | | | | | | | |
Db 2 ctcggtcacagtg 16

RESULT 8

US-09-879-341-12
; Sequence 12, Application US/09879341
; GENERAL INFORMATION:
; APPLICANT: Kster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe LLP
; STREET: 4350 La Jolla Village Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92122-1246
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/879,341
; FILING DATE: 11-Jun-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/786,416
; FILING DATE: 28-FEB-2001
; APPLICATION NUMBER: 09/287,679
; FILING DATE: 06-APR-1999
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002M
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-879-341-12

Query Match 58.0%; Score 11.6; DB 6; Length 25;
Best Local Similarity 77.8%; Pred. No. 1.8e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 gcggggtcacagtggtcg 20
| | | | | | | | | |
Db 6 GCGGCAACACAGTCGTG 23

RESULT 9

US-09-879-341-10/c
; Sequence 10, Application US/09879341
; GENERAL INFORMATION:
; APPLICANT: Kster, Hubert
; TITLE OF INVENTION: DNA Diagnostics Based on Mass Spectrometry
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe LLP
; STREET: 4350 La Jolla Village Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92122-1246
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/879,341
; FILING DATE: 11-Jun-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/786,416
; FILING DATE: 28-FEB-2001
; APPLICATION NUMBER: 09/287,679
; FILING DATE: 06-APR-1999
; APPLICATION NUMBER: 08/617,256
; FILING DATE: 18-MAR-1996
; APPLICATION NUMBER: 08/406,199
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24736-2002M
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-879-341-10

Query Match 58.0%; Score 11.6; DB 6; Length 26;
Best Local Similarity 77.8%; Pred. No. 1.8e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 3 gcgggctcacagtgtcg 20
||||| ||||| |||
Db 20 GCGGGCAACAGTCGTG 3

RESULT 10
US-09-915-060-9
; Sequence 9, Application US/09915060
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut voor Biotechnol
; TITLE OF INVENTION: Novel internal ribosome entry site, vector containing same and th
; FILE REFERENCE: 2676-4976US
; CURRENT APPLICATION NUMBER: US/09/915,060
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 99200216.2
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: in frame NotI
US-09-915-060-9

Query Match 58.0%; Score 11.6; DB 7; Length 30;
Best Local Similarity 77.8%; Pred. No. 1.9e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 cgcgggctcacagtgtcg 19
||||| ||||| |||
Db 12 cgcggccgcagagtggac 29

RESULT 11
US-09-915-060-43
; Sequence 43, Application US/09915060
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut voor Biotechnol
; TITLE OF INVENTION: Novel internal ribosome entry site, vector containing same and th
; FILE REFERENCE: 2676-4976US
; CURRENT APPLICATION NUMBER: US/09/915,060
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 99200216.2
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: primer
US-09-915-060-43

Query Match 58.0%; Score 11.6; DB 7; Length 30;
Best Local Similarity 77.8%; Pred. No. 1.9e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 cgcgggctcacagtgtcg 19
||||| ||||| |||
Db 12 cgcggccgcagagtggac 29

RESULT 12
US-60-298-340-21/c
; Sequence 21, Application US/60298340
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN GROWTH FACTOR WITH HOMOL
; FILE REFERENCE: D0166 PSP
; CURRENT APPLICATION NUMBER: US/60/298,340
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-298-340-21

Query Match 58.0%; Score 11.6; DB 8; Length 39;
Best Local Similarity 77.8%; Pred. No. 1.9e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 gcgggctcacagtgtcg 20
||||| ||||| |||
Db 25 GCTGTCTCACAGCGCGG 8

RESULT 13
US-09-787-252-17/c
; Sequence 17, Application US/09787252
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF RHO G EXPRESSION
; FILE REFERENCE: RTSP-0106
; CURRENT APPLICATION NUMBER: US/09/787,252
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/161,015
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 17
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-787-252-17

Query Match 57.0%; Score 11.4; DB 7; Length 18;
Best Local Similarity 92.3%; Pred. No. 2.3e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 cgggctcacagtg 16
||||| |||||
Db 17 CGGGCGCACAGTG 5

RESULT 14
US-09-801-274-629
; Sequence 629, Application US/09801274
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2009-001
; CURRENT APPLICATION NUMBER: US/09/801,274
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/187,510
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 60/206,129
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 1802
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 629
```

; LENGTH: 31
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-801-274-629

Query Match 57.0%; Score 11.4; DB 6; Length 31;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 gggtcacagtggctcg 20
|| ||||| :|
Db 3 gggtcacagtggcsg 17

RESULT 15
US-09-735-271-1558/C
; Sequence 1558, Application US/09735271
; GENERAL INFORMATION:
; APPLICANT: Daly, Mark J.
; APPLICANT: Hudson, Thomas J.
; APPLICANT: Lander, Eric S.
; APPLICANT: Rioux, John
; APPLICANT: Siminovitch, Kathy
; TITLE OF INVENTION: IBD-RELATED POLYMORPHISMS
; FILE REFERENCE: 2825.1025-002
; CURRENT APPLICATION NUMBER: US/09735,271
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/170,257
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 60/196,046
; PRIOR FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 2058
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1558
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(40)
; OTHER INFORMATION: n = A,T,C or G
US-09-735-271-1558

Query Match 57.0%; Score 11.4; DB 7; Length 40;
Best Local Similarity 92.3%; Pred. No. 2.4e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 qcggggtcacagt 15
||||| :|||
Db 35 GCGGGGTACAGT 23

Search completed: October 2, 2001, 16:54:58
Job time: 17661 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:47 ; Search time 417.38 Seconds
(without alignments)
9.071 Million cell updates/sec

Title: US-09-757-100B-9

Perfect score: 20

Sequence: 1 ttttctagatgctaggtat 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	ID	Description
1	20	100.0	20	3	US-09-377-310-9
2	15	75.0	15	3	US-09-377-310-29
3	15	75.0	20	3	US-09-377-310-8
4	13.6	68.0	31	1	US-08-219-012-39
5	13.6	68.0	31	4	US-08-687-421-227
6	13	65.0	15	3	US-09-377-310-28
7	12.2	61.0	29	3	US-08-792-832A-15
8	12.2	61.0	29	3	US-09-033-055A-10
9	12.2	61.0	35	4	US-08-686-968C-28
10	12.2	61.0	35	4	US-08-686-968C-30
11	12.2	61.0	35	4	US-08-686-968C-32
12	12.2	61.0	35	4	US-08-686-968C-34
13	12.2	61.0	41	2	US-08-478-386A-8
14	12.2	61.0	41	2	US-08-292-597-8
15	12.2	61.0	41	2	US-08-388-653-8
16	12.2	61.0	41	2	US-08-473-985-8
17	12.2	61.0	41	2	US-08-483-898-8
18	12.2	61.0	41	3	US-09-087-716-8
19	12.2	61.0	41	3	US-09-157-753-8
20	12.2	61.0	41	3	US-09-157-230-8
21	12.2	61.0	41	3	US-09-087-811-8
22	12.2	61.0	41	3	US-09-156-855-8
23	12.2	61.0	41	3	US-09-158-010-8
24	12.2	61.0	41	4	US-09-087-647-8
25	12	60.0	35	1	US-08-810-116-2
26	12	60.0	35	2	US-07-930-548A-2
27	12	60.0	35	2	US-08-818-604-22

c

ALIGNMENTS

RESULT 1

US-09-377-310-9
; Sequence 9, Application US/09377310B
; Patent No. 6133031

GENERAL INFORMATION:

; APPLICANT: Monia, Brett P.

; APPLICANT: Gaarde, William A.

; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase

; TITLE OF INVENTION: Expression

; FILE REFERENCE: ISPH-0389

; CURRENT APPLICATION NUMBER: US/09/377,310B

; CURRENT FILING DATE: 1999-08-19

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; LENGTH: 20

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: antisense sequence

US-09-377-310-9

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttttctagatgctaggtat 20

Db 1 ttttctagatgctaggtat 20

RESULT 2

US-09-377-310-29

; Sequence 29, Application US/09377310B

; Patent No. 6133031

GENERAL INFORMATION:

; APPLICANT: Monia, Brett P.

; APPLICANT: Gaarde, William A.

; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase

; TITLE OF INVENTION: Expression

; FILE REFERENCE: ISPH-0389

; CURRENT APPLICATION NUMBER: US/09/377,310B

; CURRENT FILING DATE: 1999-08-19

; NUMBER OF SEQ ID NOS: 43

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 29

; LENGTH: 15

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-29

Query Match 75.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ttgctagatgctagg 17
Db 1 ttgctagatgctagg 15

RESULT 3
US-09-377-310-8
; Sequence 8, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-8

Query Match 75.0%; Score 15; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctagatgctaggat 20
Db 1 ctagatgctaggat 15

RESULT 4
US-08-219-012-39
; Sequence 39, Application US/08219012
; Patent No. 5543293
; GENERAL INFORMATION:
; APPLICANT: Larry Gold
; APPLICANT: Diane Tasset
; TITLE OF INVENTION: Ligands of Thrombin
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beaton & Swanson, P.C.
; STREET: 4582 South Ulster Street Parkway, Suite #
; STREET: 403
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80237
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,012
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: none

; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: (303) 850-9900
; TELEFAX: (303) 850-9401
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-219-012-39

Query Match 68.0%; Score 13.6; DB 1; Length 31;
Best Local Similarity 80.0%; Pred. No. 98;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 tttagctagatgctaggat 20
Db 2 TTTTGGTATAGCTAGGTGT 21

RESULT 5
US-08-687-421-227
; Sequence 227, Application US/08687421
; Patent No. 617557
; GENERAL INFORMATION:
; APPLICANT: Gold, Larry
; APPLICANT: Janjic, Nebojsa
; APPLICANT: Tasset, Diane
; TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF BASIC
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR AND
; TITLE OF INVENTION: THROMBIN
; NUMBER OF SEQUENCES: 445
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,421
; FILING DATE: 08-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,005
; FILING DATE: 10-FEBRUARY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 22-APRIL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/219,012
; FILING DATE: 28-MARCH-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,333
; FILING DATE: 11-NOVEMBER-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson

REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER: NEX07/PCT
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 227:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-687-421-227

Query Match 68.0%; Score 13.6; DB 4; Length 31;
Best Local Similarity 80.0%; Pred. No. 98;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttttgctagatgctaggtat 20
||||| ||| ||||| |||
Db 2 TTTTGGTATAGGCTAGGTGT 21

RESULT 6

US-09-377-310-28
Sequence 28, Application US/09377310B
Patent No. 6133031
GENERAL INFORMATION:

APPLICANT: Monia, Brett P.
APPLICANT: Gaarde, William A.
TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
FILE REFERENCE: ISPH-0389
CURRENT APPLICATION NUMBER: US/09/377,310B
CURRENT FILING DATE: 1999-08-19
NUMBER OF SEQ ID NOS: 43
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 28
LENGTH: 15

TYPE: DNA
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: antisense sequence
US-09-377-310-28

Query Match 65.0%; Score 13; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 agatgctaggtat 20
||||| ||||| ||||| |||
Db 1 agatgctaggtat 13

RESULT 7

US-08-792-832A-15
Sequence 15, Application US/08792832A
Patent No. 6017734
GENERAL INFORMATION:

APPLICANT: Summers Dr., Max D.
APPLICANT: Braunagel Dr., Sharon C.
APPLICANT: Hong Dr., Tao
TITLE OF INVENTION: UNIQUE NUCLEOTIDE AND AMINO ACID
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/792,832A
FILING DATE: 30-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/000,955
FILING DATE: 07-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/678,435
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hibler, David W.
REGISTRATION NUMBER: 41,071
REFERENCE/DOCKET NUMBER: TAMK:190
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-792-832A-15

Query Match 61.0%; Score 12.2; DB 3; Length 29;
Best Local Similarity 82.4%; Pred. No. 5.3e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ttgctagatgctaggtat 19
||||| ||| ||||| |||
Db 13 TTGCTATTGTATAGGTA 29

RESULT 8

US-09-033-055A-10
Sequence 10, Application US/09033055A
Patent No. 6069241
GENERAL INFORMATION:

APPLICANT: OHKAWA, HIDEO
APPLICANT: IMAISHI, HIROMASA
TITLE OF INVENTION: NOVEL CYTOCHROME P450 GENE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,055A
FILING DATE: MARCH 2, 1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PAUL E. WHITE, JR.
REGISTRATION NUMBER: 32,011
REFERENCE/DOCKET NUMBER: 9437/251563
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3651
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 29

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-09-033-055A-10

Query Match 61.0%; Score 12.2; DB 3; Length 29;

Best Local Similarity 82.4%; Pred. No. 5.3e+02;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ttgctagatgctaggta 19

||||| ||||| |||||

Db 2 TTGCTTATTGCTAGGTA 18

RESULT 9

US-08-686-968C-28

; Sequence 28, Application US/08686968C

; Patent No. 6221361

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D.

; APPLICANT: Junker, David E.

; TITLE OF INVENTION: Recombinant Swinepox Virus

; FILE REFERENCE: 39119-H/JML

; CURRENT APPLICATION NUMBER: US/08/686,968C

; CURRENT FILING DATE: 1996-07-25

; NUMBER OF SEQ ID NOS: 231

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 28

; LENGTH: 35

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: primer

US-08-686-968C-28

Query Match

Best Local Similarity 82.4%; Pred. No. 5.4e+02;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ttgctagatgctaggtt 18

||||| ||||| |||||

Db 18 ttcttagtgctagtt 34

RESULT 10

US-08-686-968C-30

; Sequence 30, Application US/08686968C

; Patent No. 6221361

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D.

; APPLICANT: Junker, David E.

; TITLE OF INVENTION: Recombinant Swinepox Virus

; FILE REFERENCE: 39119-H/JML

; CURRENT APPLICATION NUMBER: US/08/686,968C

; CURRENT FILING DATE: 1996-07-25

; NUMBER OF SEQ ID NOS: 231

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 30

; LENGTH: 35

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: primer

US-08-686-968C-30

Query Match

Best Local Similarity 82.4%; Pred. No. 5.4e+02;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ttgctagatgctaggtt 18

||||| ||||| |||||

Db 18 ttcttagtgctagtt 34

RESULT 11

US-08-686-968C-32

; Sequence 32, Application US/08686968C

; Patent No. 6221361

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D.

; APPLICANT: Junker, David E.

; TITLE OF INVENTION: Recombinant Swinepox Virus

; FILE REFERENCE: 39119-H/JML

; CURRENT APPLICATION NUMBER: US/08/686,968C

; CURRENT FILING DATE: 1996-07-25

; NUMBER OF SEQ ID NOS: 231

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 32

; LENGTH: 35

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: primer

US-08-686-968C-32

Query Match

Best Local Similarity 82.4%; Pred. No. 5.4e+02;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ttgctagatgctaggtt 18

||||| ||||| |||||

Db 18 ttcttagtgctagtt 34

RESULT 12

US-08-686-968C-34

; Sequence 34, Application US/08686968C

; Patent No. 6221361

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D.

; APPLICANT: Junker, David E.

; TITLE OF INVENTION: Recombinant Swinepox Virus

; FILE REFERENCE: 39119-H/JML

; CURRENT APPLICATION NUMBER: US/08/686,968C

; CURRENT FILING DATE: 1996-07-25

; NUMBER OF SEQ ID NOS: 231

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 34

; LENGTH: 35

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: primer

US-08-686-968C-34

Query Match

Best Local Similarity 82.4%; Pred. No. 5.4e+02;

Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 ttgctagatgctaggtt 18

||||| ||||| |||||

Db 18 ttcttagtgctagtt 34

RESULT 13

US-08-478-386A-8

; Sequence 8, Application US/08478386A

; Patent No. 5830462

GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.
APPLICANT: Schreiber, Stuart L.
APPLICANT: Wandless, Thomas J.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Peter
APPLICANT: Belshaw, Thomas J.
TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
GENES AND OTHER BIOLOGICAL EVENTS
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARIAD Pharmaceuticals, Inc.
STREET: 26 Landsdowne Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC/DOS/MS/DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,386A
FILING DATE: 07/JUN/1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Figg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 2054-114A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6..11
OTHER INFORMATION: /note= "Xho I restriction site."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12..41
OTHER INFORMATION: /note= "Region of homology with
target sequence."
FEATURE:
NAME/KEY: CDS
LOCATION: 9..41
FEATURE:
NAME/KEY: misc_feature
LOCATION: 28
OTHER INFORMATION: /note= "A to G."
US-08-478-386A-8

Query Match 61.0%; Score 12.2; DB 2; Length 41;
Best Local Similarity 82.4%; Pred. No. 5.5e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Caps 0;

QY 2 ttgctagatgctaggt 18
| ||||| |||||
Db 13 TCTGCTACTTGCTAGGT 29

RESULT 14
US-08-292-597-8
Sequence 8, Application US/08292597
Patent No. 5834266
GENERAL INFORMATION:
APPLICANT: Gerald R. Crabtree

APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
TITLE OF INVENTION: Regulated Apoptosis
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARIAD Pharmaceuticals, Inc.
STREET: 26 Landsdowne Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC/DOS/MS/DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/292,597
FILING DATE: 18/AUG/1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Figg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 2054-108A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6..11
OTHER INFORMATION: /note= "Xho I restriction site."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12..41
OTHER INFORMATION: /note= "Region of homology with
target sequence."
FEATURE:
NAME/KEY: CDS
LOCATION: 9..41
FEATURE:
NAME/KEY: misc_feature
LOCATION: 28
OTHER INFORMATION: /note= "A to G."
US-08-292-597-8

Query Match 61.0%; Score 12.2; DB 2; Length 41;
Best Local Similarity 82.4%; Pred. No. 5.5e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Caps 0;

QY 2 ttgctagatgctaggt 18
| ||||| |||||
Db 13 TCTGCTACTTGCTAGGT 29

RESULT 15
US-08-388-653-8
Sequence 8, Application US/08388653
Patent No. 5869337
GENERAL INFORMATION:
APPLICANT: Crabtree, Gerald R.

APPLICANT: Schreiber, Stuart L.
APPLICANT: Spencer, David M.
APPLICANT: Wandless, Thomas J.
APPLICANT: Belshaw, Peter
TITLE OF INVENTION: REGULATED TRANSCRIPTION OF TARGETED
TITLE OF INVENTION: GENES AND OTHER BIOLOGICAL EVENTS
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: ARIAD Pharmaceuticals, Inc.
STREET: 26 Landsdowne Street
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02139
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC/DOS/MS/DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,653
FILING DATE: 14-FEB-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,386
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Figg, E. Anthony
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 2054-114A
TELEPHONE: (202) 783-6040
TELEFAX: (202) 783-6031
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6..11
OTHER INFORMATION: /note= "Xho I restriction site."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 12..41
OTHER INFORMATION: /note= "Region of homology with
OTHER INFORMATION: target sequence."
FEATURE:
NAME/KEY: CDS
LOCATION: 9..41
FEATURE:
NAME/KEY: misc_feature
LOCATION: 28
OTHER INFORMATION: /note= "A to G."
US-08-388-653-8

Query Match 61.0%; Score 12.2; DB 2; Length 41;
Best Local Similarity 82.4%; Pred. No. 5.5e+02;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 ttgtctagatgctaggt 18
| | | | | | | | | |
Db 13 TCTGCTACTTGCTAGGT 29

Search completed: October 2, 2001, 16:03:48
Job time: 14592 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:38 ; Search time 876.95 Seconds
(without alignments)
14.320 Million cell updates/sec

Title: US-09-757-100B-9

Perfect score: 20
Sequence: 1 tttagctagctagctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601:*

1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT:*

2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT:*

3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT:*

4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT:*

5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT:*

6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT:*

7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT:*

8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT:*

9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT:*

10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT:*

11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT:*

12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT:*

13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT:*

14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT:*

15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT:*

16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT:*

17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT:*

18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:*

19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT:*

20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:*

21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT:*

22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	22	Human focal adhesi
2	15	75.0	15	22	Human focal adhesi
3	15	75.0	20	22	Human focal adhesi
4	13.8	69.0	20	20	PEBP2 alpha A gene
5	13.8	69.0	36	21	HIV gp140 gene PCR
6	13.6	68.0	31	16	Thrombin DNA ligand
7	13.6	68.0	31	22	Thrombin high affi
8	13.2	66.0	27	21	CaMv Calibrator pr
9	13.2	66.0	31	21	Human genomic DNA
10	13	65.0	15	22	Human focal adhesi
11	12.6	63.0	36	21	Nucleotide sequenc

c	12	12.6	63.0	36	21	AAZ99289	Nucleotide sequenc
	13	12.4	62.0	27	20	AAV99615	Maize rpoB gene pr
	14	12.4	62.0	27	21	AAA48985	Mutagenic PCR prim
	15	12.2	61.0	29	18	AAV59549	Reverse primer amp
	16	12.2	61.0	29	19	AAV44905	Petunia flower bud
	17	12.2	61.0	35	19	AAV26145	CPV VP2 gene PCR p
	18	12.2	61.0	35	19	AAV26147	CPV VP2 gene PCR p
	19	12.2	61.0	35	19	AAV26149	CPV VP1/2 gene PCR
	20	12.2	61.0	35	19	AAV26151	CPV VP1/2 gene PCR
c	21	12.2	61.0	38	14	AA038730	Primer OMB22. Syn
	22	12.2	61.0	41	17	AA017092	CD3 transmembrane
	23	12.2	61.0	41	19	AAV62947	Murine signalling
	24	12.2	61.0	41	20	AAV02031	Murine signalling
	25	12.2	61.0	41	21	AAA60480	Murine IL2 recepto
	26	12.2	61.0	41	21	AAA14290	Murine IL2 recepto
	27	12.2	61.0	41	21	AAZ86759	PCR primer #6050 f
	28	12.2	61.0	41	22	AAZ27538	Primer #3. Uniden
	29	12.2	61.0	41	22	AAZ83458	Primer #3 used in
	30	12	60.0	26	19	AAV65088	Human ZFP PCR prim
c	31	12	60.0	27	20	AAZ25823	HPV16 E6-coding re
c	32	12	60.0	29	20	AAV05156	3' junction sequen
	33	12	60.0	31	19	AAV67544	Nucleotide fragmen
	34	12	60.0	33	20	AAZ89599	PCR primer for amp
	35	12	60.0	35	19	AAV34760	Human KDR primer 2
c	36	12	60.0	39	16	AAQ84442	Mycobacterium fort
c	37	12	60.0	42	16	AAT01098	V1Jns-gp160IIIB/IR
c	38	12	60.0	43	16	AAT01135	HIV IIIB gp160 PCR
c	39	12	60.0	43	16	AAT01136	HIV IIIB gp160 PCR
c	40	12	60.0	43	18	AAZ85524	HIV IIIB gp160 PCR
c	41	12	60.0	43	18	AAZ85525	HIV IIIB gp160 PCR
c	42	12	60.0	43	19	AAV21751	HIV IIIB gp160 gen
c	43	12	60.0	43	19	AAV21752	HIV IIIB gp160 gen
c	44	11.8	59.0	20	14	AA041648	PKPIAV expression
	45	11.8	59.0	22	18	AAZ83010	Primer C to isolat

ALIGNMENTS

RESULT 1

AAZ5541

ID AAC65541 standard; DNA; 20 BP.

XX

AC AAC65541;

XX

DT 12-FEB-2001 (first entry)

XX

XX Human focal adhesion kinase antisense sequence #7.

DE Human; focal adhesion kinase; FAK; signal transduction; cancer;

XX embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphorothioate; ss.

XX

OS Homo sapiens.

XX

XX US6133031-A.

XX

PD 17-OCT-2000.

XX

XX 19-AUG-1999; 99US-0377310.

PF

XX 19-AUG-1999; 99US-0377310.

PR

XX (ISIS-) ISIS PHARM INC.

XX

XX Monia BP, Gaarde WA;

XX

XX WPI; 2001-006141/01.

XX

XX New antisense compounds for inhibiting focal adhesion kinase

PT expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

XX

PS Claim 3; Column 23; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX Sequence 20 BP; 4 A; 2 C; 5 G; 9 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttctgtagatgctaggtat 20

Db 1 ttctgtagatgctaggtat 20

RESULT 2

AAC65561
 ID AAC65561 standard; DNA; 15 BP.

AC AAC65561;

DT 12-FEB-2001 (first entry)

XX Human focal adhesion kinase antisense sequence #27.

XX Human; focal adhesion kinase; FAK; signal transduction; cancer;

KW embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphorothioate; ss.

OS Homo sapiens.

XX US6133031-A.

PD 17-OCT-2000.

XX 19-AUG-1999; 99US-0377310.

PR 19-AUG-1999; 99US-0377310.

XX (ISIS-) ISIS PHARM INC.

PI Monia BP, Gaarde WA;

XX WPI; 2001-006141/01.

XX New antisense compounds for inhibiting focal adhesion kinase

PT expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

PS Example 2; Column 25; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX Sequence 15 BP; 3 A; 2 C; 5 G; 5 T; 0 other;

Query Match

Best Local Similarity 75.0%; Score 15; DB 22; Length 15;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ttgctagatgctagg 17

Db 1 ttgctagatgctagg 15

RESULT 3

AAC65540

ID AAC65540 standard; DNA; 20 BP.

XX AAC65540;

DT 12-FEB-2001 (first entry)

XX Human focal adhesion kinase antisense sequence #6.

XX Human; focal adhesion kinase; FAK; signal transduction; cancer;

KW embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphorothioate; ss.

OS Homo sapiens.

XX US6133031-A.

PD 17-OCT-2000.

PF 19-AUG-1999; 99US-0377310.

PR 19-AUG-1999; 99US-0377310.

XX (ISIS-) ISIS PHARM INC.

PI Monia BP, Gaarde WA;

XX WPI; 2001-006141/01.

XX New antisense compounds for inhibiting focal adhesion kinase

PT expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

XX Claim 3; Column 23; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX Sequence 20 BP; 4 A; 4 C; 5 G; 7 T; 0 other;

Query Match 75.0%; Score 15; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 63;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ctatagatgctaggtat 20

Db 1 ctatagatgctaggtat 15

RESULT 4

AAX33269

ID AAX33269 standard; DNA; 20 BP.

XX AAX33269;

DT 30-JUN-1999 (first entry)

XX PEP2 alpha A gene expression regulating DNA PCR primer SEQ ID NO:26.

XX PEP2 alpha A gene; expression; regulation; bone disease;

KW osteoporosis; PCR primer; ss.

XX OS Synthetic.
 XX PN WO9911787-A1.
 XX PD 11-MAR-1999.
 XX PF 02-SEP-1998; 98WO-JP03920.
 XX PR 08-APR-1998; 98JP-0114135.
 XX PR 02-SEP-1997; 97JP-0254250.
 XX PR 15-OCT-1997; 97JP-0299407.
 XX PA (SUMU) SUMITOMO PHARM CO LTD.
 XX PI Fujiwara M, Harada H, Katsumata T, Nakatsuka M;
 XX PI Ogawa S, Tagashira S;
 XX DR WPI; 1999-243621/20.
 XX CC DNA regulating expression of PEBP2 alphaA gene to produce regulator
 PT protein, useful as promoter for prevention or/and treatment of bone
 PT diseases e.g. osteoporosis
 XX PS Example 6; Page 39; 118pp; Japanese.
 XX CC The present invention describes DNA which participates in the regulation
 CC of expression of PEBP2 alpha A gene. The DNA produces a regulator
 CC protein with the activity of promoting bone formation and can serve as a
 CC promoter for prevention and treatment of bone diseases including
 CC osteoporosis. The present sequence represents a PCR primer used in an
 CC example from the present invention.
 XX CC Sequence 20 BP; 6 A; 1 C; 6 G; 7 T; 0 other;
 SQ
 Query Match 69.0%; Score 13.8; DB 20; Length 20;
 Best Local Similarity 88.2%; Pred. No. 2.6e+02;
 . Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 ttgtctagatgctaggt 18
 ||||| || |||||
 Db 3 ttgtataaatgctaggt 19
 RESULT 5
 AAA48327/C
 ID AAA48327 standard; DNA; 36 BP.
 AC AAA48327;
 XX 28-SEP-2000 (first entry)
 DE HIV gp140 gene PCR primer, gp140CysEcoRI.
 XX KW Antigen presentation; vaccine; infectious disease; allergy; cancer;
 KW molecular scaffold; immune response; farm animal; organiser; gp140;
 KW immunostimulatory; cytostatic; antiallergy; human immunodeficiency virus;
 KW HIV; PCR primer; ss.
 XX OS Human immunodeficiency virus.
 XX PN WO200032227-A2.
 XX PD 08-JUN-2000.
 XX PF 30-NOV-1999; 99WO-IB01925.
 XX PR 30-NOV-1998; 98US-0110414.
 XX PR 08-JUL-1999; 99US-0142788.
 XX PA (CYTO-) CYTOS BIOTECHNOLOGY AG.
 XX

PI Renner WA, Hennecke F, Nieba L, Bachmann M;
 DR WPI; 2000-412159/35.
 XX Composition for use as vaccine against infectious diseases and in
 PT treatment of cancer and allergies comprises non-naturally occurring
 PT molecular scaffold and antigen or antigenic determinant -
 XX Example 26; Page 77; 102pp; English.
 XX CC A new method for developing vaccines has been identified, in which a
 CC non-naturally occurring molecular scaffold, having a core particle and a
 CC covalently attached organiser, is attached to an antigen or antigenic
 CC determinant. The scaffold and antigen or antigenic determinant interact
 CC to form an ordered and repetitive antigen array. The composition is
 CC useful as a vaccine against infectious diseases, to induce immune
 CC responses in farm animals and also in the treatment of cancer and
 CC allergies. The present sequence is the PCR primer, gp140CysEcoRI. This
 CC primer was used to amplify the coding sequence for the gp140 gene of
 CC human immunodeficiency virus (HIV). The PCR product was then used as an
 CC antigen in the present invention.
 XX CC Sequence 36 BP; 11 A; 10 C; 6 G; 9 T; 0 other;
 SQ
 Query Match 69.0%; Score 13.8; DB 21; Length 36;
 Best Local Similarity 88.2%; Pred. No. 2.8e+02;
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4 tgctagatgctaggtat 20
 ||||| ||||| ||
 Db 22 TGCTAGCTGCTAGGAAT 6
 RESULT 6
 AAT00213
 ID AAT00213 standard; DNA; 31 BP.
 XX AC AAT00213;
 XX 14-AUG-1996 (first entry)
 DT Thrombin DNA ligand, clone #17.
 DE Family 1; family 2; ligand: thrombin;
 KW systematic evolution of ligands by exponential enrichment; SELEX;
 KW heparin; selection; region of homology; inhibitor; ss.
 XX OS Synthetic.
 XX PN WO9521853-A1.
 XX PD 17-AUG-1995.
 XX PF 06-FEB-1995; 95WO-US01458.
 XX PR 28-MAR-1994; 94US-0219012.
 XX PR 10-FEB-1994; 94US-0195005.
 XX PR 11-JUN-1990; 90US-0536428.
 XX PR 10-JUN-1991; 91US-0714131.
 XX PR 22-APR-1993; 93US-0061691.
 XX PA (NEXS-) NEXSTAR PHARM INC.
 XX PI Gold L, Janjic N, Tasset D;
 XX DR WPI; 1995-293073/38.
 XX PT Identification of ligands to basic fibroblast growth factor and
 PT thrombin - which can be modified for increased in vivo stability
 XX Claim 39; Page 95; 236pp; English.
 XX

CC The sequences given in AAT00202-25 and AAT00227-57 represent two groups
 CC of ligands to thrombin. These sequences were isolated using the single
 CC stranded DNA molecules given in AAT00201 and AAT00226 which comprise a
 CC 30N and a 60N variable region, respectively. These ligands were
 CC isolated using systematic evolution of ligands by exponential enrichment
 CC (SELEX). The selection was conducted in a buffer solution at 37 deg. C.
 CC After 12 rounds of selection, no additional improvement in binding was
 CC seen. By studying regions of homology between the isolated ligands, a
 CC truncated ligand of 38 nucleotides (see AA098403-04) was identified which
 CC retains high affinity binding and inhibits clotting. These ligands are
 CC inhibitors of thrombin and are therefore useful in treating thrombin
 CC mediated conditions and in studying the structure and binding of
 CC thrombin.

XX Sequence 31 BP; 5 A; 3 C; 11 G; 12 T; 0 other;

Query Match 68.0%; Score 13.6; DB 16; Length 31;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ttctgtagctagctaggtat 20
 Db 2 ttctgtagctagctaggtat 21

RESULT 7
 AAF70765
 ID AAF70765 standard; DNA; 31 BP.

XX AC AAF70765;

XX DT 20-APR-2001 (first entry)

XX DE Thrombin high affinity ligand #12.

XX KW Ligand; basic fibroblast growth factor; bFGF; gene therapy; vascular;
 XX KW atherosclerosis; angioplasty; stability; ss.

XX OS Unidentified.

XX PN US6177557-B1.

XX PD 23-JAN-2001.

XX PF 05-AUG-1996; 96US-0687421.

XX PR 11-JUN-1990; 90US-0536428.

XX PR 10-JUN-1991; 91US-0714131.

XX PR 06-NOV-1992; 92US-0973333.

XX PR 10-FEB-1994; 94US-0195005.

XX PR 28-MAR-1994; 94US-0219012.

XX PA (NEXS-) NEXSTAR PHARM INC.

XX PI Janjic N, Gold L, Tasset D;

XX DR WPI; 2001-158583/16.

XX PT Novel nucleic acid ligands to basic fibroblast growth factor that are
 XX PT useful as inhibitors of basic fibroblast growth factors and 2'-amino
 XX PT modified RNA ligands, exhibit increased in vivo stability -

XX PS Example 19; Column 57-58; 153pp; English.

XX CC The present invention relates to a purified and isolated non-naturally
 XX CC occurring DNA ligands to basic fibroblast growth factor (bFGF).

XX CC The ligands are useful as part of gene therapy treatments and
 XX CC for diagnosing pathogenesis of vascular diseases including

XX CC initiation and progression of atherosclerosis, acute coronary
 XX CC syndromes, vein graft disease and restenosis following coronary

XX CC angioplasty. The ligands have improved stability in vivo.

SQ Sequence 31 BP; 5 A; 3 C; 11 G; 12 T; 0 other;

Query Match 68.0%; Score 13.6; DB 22; Length 31;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ttctgtagctagctaggtat 20
 Db 2 ttctgtagctagctaggtat 21

RESULT 8

AAZ93950

ID AAZ93950 standard; DNA; 27 BP.

XX AC AAZ93950;

XX DT 29-AUG-2000 (first entry)

XX DE CaMV Calibrator probe sequence used in detection method.

XX KW Detection; identification; PCR; polymerase chain reaction; probe;
 XX KW primer; calibrator; diagnosis; pathogen; virus; bacteria; HHV-6;
 XX KW HHV-7; HHV-8; human herpes virus; human immunodeficiency virus; HIV;
 XX KW AIDS; acquired immune deficiency syndrome;
 XX KW Mycobacterium tuberculosis; ss.

XX OS Synthetic.

XX PN WO200029613-A1.

XX PD 25-MAY-2000.

XX PF 17-NOV-1999; 99WO-EP08847.

XX PR 17-NOV-1998; 98IT-MI02491.

XX PA (SANR-) FOND CENT SAN RAFFAELE DEL MONTE TABOR.

XX PI Locatelli G, Lusso P, Malnati M, Salvatori F, Scarlatti G;

XX DR WPI; 2000-387819/33.

XX PT Quantitative detection of nucleic acids, based on the use of a
 XX PT calibrator, suitable primers and probes, and a nucleic acid polymerase
 XX PT with 5'-3' nuclease activity

XX PS Example 1; Page 13; 39pp; English.

XX CC A new method for the quantitative detection of nucleic acids in a
 XX CC sample which uses a calibrator, suitable primers and probes, and
 XX CC a nucleic acid polymerase with 5'-3' nuclease activity is described.
 XX CC The method comprises adding a calibrator nucleotide to the sample,
 XX CC the calibrator nucleotide having the same sequence as the target with
 XX CC the exception of one or more regions which in the target hybridize
 XX CC with a probe labeled with a reporter and quencher, or which hybridize
 XX CC with the probe and with 2 or more primers. These regions have a
 XX CC different, randomized nucleotide sequence and a similar Tm. Then
 XX CC extracting the calibrator and target nucleotides from the sample.
 XX CC Probes and primers are then added to the extracted sample/calibrator
 XX CC mixture and PCR is performed. Reaction in the presence of the target
 XX CC nucleic acid specific probe permits quantitation of the copy number
 XX CC of the extracted target nucleic acid. The reaction in the presence of
 XX CC the calibrator nucleotide permits quantitation of the calibrator copy
 XX CC number. The reaction in the presence of both permits calculation of
 XX CC the total number of target templates and calibrator, allowing
 XX CC calculation of the percentage of calibrator recovery. The method may
 XX CC be used for the diagnosis of viral and/or other pathogenic agents in
 XX CC body fluids, and to monitor safety and/or genetic composition of
 XX CC waters, foods, and plant species used in the alimentary field. The
 XX CC method is used to quantitatively detect the genomic nucleic acid of
 XX CC human herpes virus (HHV)-6, HHV-7, HHV-8 and human immunodeficiency

CC virus (HIV). The method provides an enhanced sensitivity, accuracy,
 CC and precision and a reduced measure viability. For specific probe,
 CC primer and calibrator sequences used in the method, see GENESEQ
 CC records AA293935-293965.

XX SQ Sequence 27 BP; 5 A; 7 C; 7 G; 8 T; 0 other;
 Query Match 66.0%; Score 13.2; DB 21; Length 27;
 Best Local Similarity 83.3%; Pred. No. 5.6e+02;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 ttgctagatgctaggtat 20
 | | | | | | | | | | | |
 Db 2 tcgctacatgctaggtat 19

RESULT 9
 AAA79014/c
 ID AAA79014 standard; DNA; 31 BP.

XX AC AAA79014;

XX DT 20-NOV-2000 (first entry)

XX DE Human genomic DNA polymorphic site sequence tag SEQ ID NO:384.

XX KW Human; genomic DNA; polymorphism; genome; allele-specific; primer;
 KW probe; hybridisation; polymorphic site; forensic; paternity testing;
 KW medicine; phenotypic trait; genetic analysis; genetic mapping; ds.

XX OS Homo sapiens.

XX PN EP1024200-A2.

XX PD 02-AUG-2000.

XX PF 26-JAN-2000; 2000EP-0250023.

XX PR 27-JAN-1999; 99US-0238402.

XX PA (AFFY-) AFFYMETRIX INC.

XX PI Patil N, Shah N, Warrington JA;

XX PS WPI; 2000-500198/45.

XX PT Human genomic polymorphic nucleic acid segments, allele specific
 PT primers and probes, and methods of analysis, useful for e.g. forensics,
 PT paternity testing, genetic mapping, -

XX PS Claim 1; Page 15; 14pp; English.

CC The present invention describes a nucleic acid segment of 10-100
 CC contiguous bases chosen from one of 632 fragments (AAA78631 to
 CC AAA79262), where the segment comprises a polymorphic site or an
 CC immediately adjacent base, or the complement of the segment. Also
 CC described are: (1) an allele-specific oligonucleotide that hybridises to
 CC a segment of the novelty; (2) an isolated nucleic acid comprising a
 CC sequence of the novelty where the polymorphic site within the sequence is
 CC occupied by a base other than the reference base indicated in the
 CC specification; and (3) analysing a nucleic acid, comprising obtaining a
 CC nucleic acid from an individual, and determining a base occupying any one
 CC of the polymorphic sites of the novelty. The nucleic acid segments and
 CC method can be used to analyse an individual's nucleic acid sequences for
 CC the presence of polymorphisms. The method can also be used to test for a
 CC disease phenotype and correlate the presence of the phenotype with a
 CC particular polymorphism. The presence of polymorphic sites are useful
 CC for, e.g. forensics, paternity testing, correlation of polymorphisms
 CC with phenotypic traits and for genetic mapping of phenotypic traits.
 CC AAA78631 to AAA79262 represent sequence tags of human genomic DNA
 CC fragments containing polymorphic sites. The base occupying the
 CC polymorphic site is indicated using IUPAC-IUB nomenclature.

XX SQ Sequence 31 BP; 9 A; 8 C; 8 G; 5 T; 1 other;

Query Match 66.0%; Score 13.2; DB 21; Length 31;
 Best Local Similarity 75.0%; Pred. No. 5.7e+02;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttgtgctagatgctaggtat 20
 | | | | | | | | | | | |
 Db 24 TTTGCTCTCATGCTAGAGAT 5

RESULT 10
 AAC65560
 ID AAC65560 standard; DNA; 15 BP.

XX AC AAC65560;

XX DT 12-FEB-2001 (first entry)

XX DE Human focal adhesion kinase antisense sequence #26.

XX KW Human; focal adhesion kinase; FAK; signal transduction; cancer;
 KW embryonic development disorder; angiogenic disorder; wound healing;
 KW antisense; phosphorothioate; ss.

XX OS Homo sapiens.

XX PN US6133031-A.

XX PD 17-OCT-2000.

XX PF 19-AUG-1999; 99US-0377310.

XX PR 19-AUG-1999; 99US-0377310.

XX PA (ISIS-) ISIS PHARM INC.

XX PI Monia BP, Gaarde WA;

XX PS WPI; 2001-006141/01.

XX PT New antisense compounds for inhibiting focal adhesion kinase
 PT expression, especially useful for inhibiting retinal
 PT neovascularization, or for diagnosing and treating e.g. colon cancer -
 XX Example 2; Column 25; 30pp; English.

XX CC The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX SQ Sequence 15 BP; 4 A; 2 C; 4 G; 5 T; 0 other;

Query Match 65.0%; Score 13; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 agatgctaggtat 20
 | | | | | | | | | | | |
 Db 1 agatgctaggtat 13

RESULT 11
 AA299288
 ID AA299288 standard; DNA; 36 BP.
 XX

```

AC AA299288;
XX
XX 03-JUL-2000 (first entry)
XX
DE Nucleotide sequence of a clone from a RNA-protein fusion library.
XX
XX RNA-protein fusion; protein-protein interaction; drug target;
KW protein chip; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 18..19
FT /*tag= a
FT /note= "an unspecified number of bases are present
FT between these nucleotides"
XX
XX WO200009737-A1.
XX
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US18603.
XX
XX 17-AUG-1998; 98US-0096818.
XX
XX (PHYL-) PHYLLOS INC.
XX
XX Hammond PW, Lipovsek D;
XX
XX WPI; 2000-224362/19.
XX
XX Producing nucleic acids lacking 3'-untranslated regions to optimise
XX production of fusion proteins, used to produce fusion libraries for
XX identification of protein:protein interactions and drug targets for the
XX production of protein chips -
XX
XX Example 1; Fig 14; 52pp; English.
XX
XX The present sequence represents a fragment from a clone of a
XX RNA-protein fusion library comprising KIAA0111. specification describes
XX a method for removing the 3'-untranslated region of a DNA molecule
XX comprising an open reading frame (ORF). The method comprises providing
XX a DNA molecule having an ORF and a 3'-untranslated region, the DNA
XX having a 5' overhang and a blunt end at the 3' end, and treating the
XX DNA molecule with a 3'-5' exonuclease followed by a single-stranded
XX nuclease to remove the 3'-untranslated region. The products can be
XX used to produce RNA-protein fusion libraries. The fusion libraries
XX can be used for the identification of protein-protein interactions,
XX identification of drug targets, and hybridisation to solid supports to
XX create protein chips (or beads). The RNA-protein molecules may be
XX arranged in spatially defined arrays on the protein chips to carry out
XX large scale screening for protein or compound identification.
XX
XX Sequence 36 BP; 6 A; 9 C; 10 G; 11 T; 0 other;
XX
Query Match 63.0%; Score 12.6; DB 21; Length 36;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ttttgctagctgctaggta 19
||||| || |||||
Db 5 ttttgatgaagctaggta 23

RESULT 12
AA299289
ID AA299289 standard; DNA; 36 BP.
XX
XX AA299289;
XX
XX 03-JUL-2000 (first entry)
XX

```

```

DE Nucleotide sequence of a clone from a RNA-protein fusion library.
XX
XX RNA-protein fusion; protein-protein interaction; drug target;
KW protein chip; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 18..19
FT /*tag= a
FT /note= "an unspecified number of bases are present
FT between these nucleotides"
XX
XX WO200009737-A1.
XX
XX 24-FEB-2000.
XX
XX 16-AUG-1999; 99WO-US18603.
XX
XX 17-AUG-1998; 98US-0096818.
XX
XX (PHYL-) PHYLLOS INC.
XX
XX Hammond PW, Lipovsek D;
XX
XX WPI; 2000-224362/19.
XX
XX Producing nucleic acids lacking 3'-untranslated regions to optimise
XX production of fusion proteins, used to produce fusion libraries for
XX identification of protein:protein interactions and drug targets for the
XX production of protein chips -
XX
XX Example 1; Fig 14; 52pp; English.
XX
XX The present sequence represents a fragment from a clone of a
XX RNA-protein fusion library comprising KIAA0111. specification describes
XX a method for removing the 3'-untranslated region of a DNA molecule
XX comprising an open reading frame (ORF). The method comprises providing
XX a DNA molecule having an ORF and a 3'-untranslated region, the DNA
XX having a 5' overhang and a blunt end at the 3' end, and treating the
XX DNA molecule with a 3'-5' exonuclease followed by a single-stranded
XX nuclease to remove the 3'-untranslated region. The products can be
XX used to produce RNA-protein fusion libraries. The fusion libraries
XX can be used for the identification of protein-protein interactions,
XX identification of drug targets, and hybridisation to solid supports to
XX create protein chips (or beads). The RNA-protein molecules may be
XX arranged in spatially defined arrays on the protein chips to carry out
XX large scale screening for protein or compound identification.
XX
XX Sequence 36 BP; 6 A; 9 C; 10 G; 11 T; 0 other;
XX
Query Match 63.0%; Score 12.6; DB 21; Length 36;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ttttgctagctgctaggta 19
||||| || |||||
Db 5 ttttgatgaagctaggta 23

RESULT 13
AAV99615/C
ID AAV99615 standard; DNA; 27 BP.
XX
XX AAV99615;
XX
XX 29-MAR-1999 (first entry)
XX
XX Maize rpoB gene primer rpoB#3.
XX
XX Promoter; nuclear encoded plastid RNA polymerase; NEP;
KW rpoB; chloroplast; transgenic plant; maize; PCR; primer; ss.

```

```

XX Synthetic.
OS Zea mays.
XX WO9855595-Al.
XX
XX 10-DEC-1998.
XX
XX 03-JUN-1998; 98WO-US11437.
XX
XX 12-SEP-1997; 97US-0058670.
XX 03-JUN-1997; 97US-0048376.
XX
XX (RUTF ) UNIV RUTGERS STATE NEW JERSEY.
XX
XX Maliga P, Silhavy D, Sritaman P;
XX WPI; 1999-070262/06.
XX
XX Isolated nuclear-encoded plastid RNA polymerase promoter sequences -
PT useful for expressing exogenous protein in plant plastids such as
PT chloroplasts
XX
XX Example 1; Page 17; 79pp; English.
XX
XX This is the nucleotide sequence of maize rpoB gene primer rpoB#3.
CC The 5' nucleotide of the primer corresponds to nucleotide 21394
CC of the complementary strand of the maize plastid genome sequence.
CC The primer was designed to add a XhoI restriction site downstream
CC of an amplified rpoB fragment following PCR amplification. The PCR
CC product was cloned into vector pBSKS+ and used to generate
CC protecting RNA for use in in vitro capping experiments. The
CC invention provides isolated rpoB, atpB, clpP and 16S rDNA NEP and
CC PEP promoter elements (see AAY9569-99) useful for producing
CC exogenous proteins of interest in plant plastids.
XX
XX Sequence 27 BP; 9 A; 5 C; 6 G; 7 T; 0 other;

Query Match 62.0%; Score 12.4; DB 20; Length 27;
Best Local Similarity 92.9%; Pred. No. 1.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ttgctagatgctag 16
Db 17 TTGCTAGATTCCTAG 4
|||||

RESULT 14
AAAA8985
ID AAA48985 standard; DNA; 27 BP.
XX
XX AAA48985;
XX
XX 28-NOV-2000 (first entry)
XX
XX Mutagenic PCR primer Bcl2-3 targeted to anti-apoptotic gene bcl-2.
XX
XX Apoptosis; bcl-2; resistance; bacterial; viral; pathogen;
KW PCR primer; human; ss.
XX
XX Homo sapiens.
XX
XX WO200026391-A2.
XX
XX 11-MAY-2000.
XX
XX 29-OCT-1999; 99WO-US25522.
XX
XX 30-OCT-1998; 98US-0106321.
XX 09-JUN-1999; 99US-0138303.
XX
XX (UYNE-) UNIV NEBRASKA-LINCOLN.
XX

```

```

XX Dickman MB;
XX
XX WPI; 2000-365634/31.
XX
XX Transgenic plants with improved resistance characteristics comprising
XX nucleic acids encoding apoptotic proteins -
XX
XX Example 3; Page 58; 109pp; English.
XX
XX The present invention relates to the use of apoptotic genes in the the
XX production of transgenic plants with improved resistance
XX characteristics. The present sequence is the mutagenic PCR primer Bcl2-3.
XX This primer was used with primer Bcl2-5 (AAA48984) to introduce a 5' NcoI
XX site and a 3' XbaI site in the human apoptotic gene bcl-2. Specifically
XX the primer introduced an Ala residue between the Met (position 1) and
XX His (position 2) residues of the native protein. The altered bcl-2
XX sequence was used in the creation of the final vector used to transform
XX plants. The improved resistance characteristics of the plants helps
XX protect against bacterial, viral and other pathogens. Resistance to
XX abiotic challenges may also be conferred.
XX
XX Sequence 27 BP; 2 A; 5 C; 8 G; 12 T; 0 other;

Query Match 62.0%; Score 12.4; DB 21; Length 27;
Best Local Similarity 92.9%; Pred. No. 1.5e+03;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ttgtctagatgct 14
Db 2 ttttctagatgct 15
|||||

RESULT 15
AAT59549
ID AAT59549 standard; DNA; 29 BP.
XX
XX AAT59549;
XX
XX 18-NOV-1997 (first entry)
XX
XX Reverse primer amplifies fragment encoding ODV-E66 residues 1-23.
XX
XX N-terminal transport polypeptide; insect pest control; cell cycle;
KW Autographa californica multinucleocapsid nuclear polyhedrosis virus;
KW AcMNPV; occlusion derived virus; ODV; 66 kD envelope protein; ODV-E66;
KW membrane protein; intranuclear viral-induced microvesicle; GFP; URF-13;
KW beta-galactosidase; b-gal; green fluorescent protein; marker; primer;
KW polymerase chain reaction; PCR; amplify; ss.
XX
XX Synthetic.
XX
XX WO9703198-A2.
XX
XX 30-JAN-1997.
XX
XX 05-JUL-1996; 96WO-US11320.
XX
XX 03-JUL-1996; 96US-0000955.
XX 07-JUL-1995; 95US-0000955.
XX
XX (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
XX Braunagel SC, Hong TC, Summers MD;
XX WPI; 1997-119053/11.
XX
XX New isolated transport polypeptide from AcMNPV - can direct
PT localisation of desired proteins to occlusion derived virus
PT envelopes
XX
XX Example 1; Page 39; 104pp; English.
XX

```

XX The sequences given in AAT59548-49 are primers which were used to
 CC amplify the sequence encoding residues 1-23 of Autographa californica
 CC multi-nucleocapsid nuclear polyhedrosis virus (ACMNPV) occlusion
 CC derived virus (ODV) 66 kb envelope protein (ODV-E66). ODV-E66 is an
 CC integral membrane protein of the ODV envelope. In an infected cell
 CC nucleus, ODV-E66 is present in the intranuclear viral-induced
 CC microvesicles and ODV envelope, providing evidence that the microvesicles
 CC function as an immediate precursor in the assembly of the ODV envelope.
 CC In infected cells, ODV-E66 localises to the ODV envelope, intranuclear
 CC microvesicles, membranes of the nuclear envelope and cytoplasmic
 CC membranes in close juxtaposition to the nuclear envelope. The amplified
 CC sequence encodes the N-terminal, target sequence domain of ODV-E66. The
 CC target sequence may be used to target beta-galactosidase (b-gal), green
 CC fluorescent protein (GFP) and URF-13 to the membranes of the nuclear
 CC envelope which are further transported into the intranuclear
 CC microvesicles and ODV envelope. The target protein can be used to target
 CC foreign proteins to the ODV envelope and intranuclear microvesicles. It
 CC can be used for insect pest control, for therapeutic applications, e.g.
 CC for location of a protein, peptide or derivative into the nuclear
 CC membrane which will alter the cell cycle and for diagnostics, e.g. to
 CC insert a marker specific for a disease or abnormality into a cell tissue
 CC manifesting such abnormalities.

XX SQ Sequence 29 BP; 5 A; 3 C; 6 G; 15 T; 0 other;

Query Match 61.0%; Score 12.2; DB 18; Length 29;
 Best Local Similarity 82.4%; Pred. No. 1.9e+03;
 Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 3 ttgctagatgctaggta 19
 ||||| |||||
 Db 13 ttgctattgataggta 29

Search completed: October 2, 2001, 16:18:38
 Job time: 15482 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:37 ; Search time 876.95 Seconds
(without alignments)
14.320 Million cell updates/sec

Title: US-09-757-100B-8

Perfect score: 20

Sequence: 1 ctatagctaggtatctgtc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

1: /SIDS8/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	22	AAC65540 Human focal adhesi
2	15.8	79.0	27	21	AAZ93950 Canv Calibrator pr
3	15	75.0	15	22	AAC65560 Human focal adhesi
4	15	75.0	20	22	AAC65541 Human focal adhesi
5	13.4	67.0	43	14	AAQ50213 c-fos position 339
6	13	65.0	38	13	AAQ28274 PCR primer 2 for a
7	12.8	64.0	24	18	AAT97310 Human Fas signal s
8	12.8	64.0	27	21	AAT97310 PCR primer for 1-2
9	12.8	64.0	39	21	AAZ37684 Human tumour suppr
10	12.6	63.0	21	16	AAT00720 Multiple tumour su
11	12.6	63.0	21	16	AAQ99928 Human MTS1E1-beta

12	12.6	63.0	21	18	AAT72309 Human multiple tum
13	12.6	63.0	21	19	AAV53849 Nucleotide sequenc
14	12.6	63.0	21	19	AAV11268 Human MTS1E1-beta
15	12.6	63.0	21	20	AAV70613 Reverse PCR primer
16	12.6	63.0	21	21	AAA95847 Human MTS1E1-beta g
17	12.6	63.0	21	21	AAA11178 Human multiple tum
18	12.6	63.0	21	21	AAZ39383 Human MTS1E1-beta
19	12.6	63.0	21	21	AAZ48786 PCR primer for hum
20	12.6	63.0	21	21	AAZ39979 PCR primer for hum
21	12.6	63.0	21	22	AAZ58201 Reverse primer #1.
22	12.6	63.0	21	22	AAC88772 Human catenin-bind
23	12.6	63.0	21	22	AAC60498 Reverse primer use
24	12.6	63.0	22	22	AAC88771 Human catenin-bind
25	12.6	63.0	37	14	AAQ49810 bGH gene 3' antise
26	12.6	63.0	39	16	AAQ85851 MHC class I allele
27	12.6	63.0	39	16	AAQ85781 MHC CW3 forward am
28	12.6	63.0	39	16	AAQ85745 MHC CW3 CDNA forwa
29	12.6	63.0	39	20	AAZ31530 Probe EF-4 for HIV
30	12.6	63.0	39	20	AAZ11016 Probe EF-7 for HIV
31	12.4	62.0	33	21	AAZ72026 HIV-1 proviral DNA
32	12.4	62.0	33	21	AAZ88080 HIV packaging cons
33	12.4	62.0	33	21	AAZ56555 pHP-1dl.28 WT nucl
34	12.4	62.0	35	19	AAV34760 Human KDR primer 2
35	12.2	61.0	18	13	AAQ34171 Downstream sequenc
36	12.2	61.0	25	21	AAZ48824 PCR primer for Hum
37	12.2	61.0	36	21	AAZ99288 Nucleotide sequenc
38	12.2	61.0	36	21	AAZ99289 Nucleotide sequenc
39	12.2	61.0	42	18	AAU86866 plasmid pBS-HIV ol
40	12.2	61.0	42	20	AAV64682 HIV anti-viral oli
41	12	60.0	15	22	AAC65561 Human focal adhesi
42	12	60.0	17	21	AAF03205 Hammerhead ribozym
43	12	60.0	17	21	AAF03206 Hammerhead ribozym
44	12	60.0	24	19	AAV15580 Primer for HIV RNA
45	12	60.0	25	22	AAF24692 PCR primer used to

ALIGNMENTS

RESULT 1
AAC65540
ID AAC65540 standard; DNA; 20 BP.
XX AC AAC65540;
XX XX
DT 12-FEB-2001 (first entry)
XX XX
XX Human focal adhesion kinase antisense sequence #6.
DE XX
DE Human; focal adhesion kinase; FAK; signal transduction; cancer;
KW embryonic development disorder; angiogenic disorder; wound healing;
KW antisense; phosphorothioate; ss.
XX OS Homo sapiens.
XX XX
XX US6133031-A.
XX XX
PD 17-OCT-2000.
XX XX
XX 19-AUG-1999; 99US-0377310.
XX XX
XX 19-AUG-1999; 99US-0377310.
XX XX
XX (ISIS-) ISIS PHARM INC.
XX PA
XX Monia BP, Gaarde WA;
XX PI
XX c-fos position 339
XX DR
XX PCR primer 2 for a
XX Human Fas signal s
XX PCR primer for 1-2
XX Human tumour suppr
XX Multiple tumour su
XX Human MTS1E1-beta

PS Claim 3; Column 23; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX Sequence 20 BP; 4 A; 4 C; 5 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.18;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ctatagtgctaggtatctgtc 20

|||||

Db 1 ctatagtgctaggtatctgtc 20

RESULT 2

AAZ93950

ID AAZ93950 standard; DNA; 27 BP.

XX AC AAZ93950;

XX DT 29-AUG-2000 (first entry)

XX DE CamV Calibrator probe sequence used in detection method.

XX KW Detection; identification; PCR; polymerase chain reaction; probe;

XX KW primer; calibrator; diagnosis; pathogen; virus; bacteria; HHV-6;

XX KW HHV-7; HHV-8; human herpes virus; human immunodeficiency virus; HIV;

XX KW AIDS; acquired immune deficiency syndrome;

XX KW Mycobacterium tuberculosis; ss.

XX OS Synthetic.

XX PN WO200029613-A1.

XX PD 25-MAY-2000.

XX PF 17-NOV-1999; 99WO-EP08847.

XX PR 17-NOV-1998; 98IT-MI02491.

XX PA (SANR-) FOND CENT SAN RAFFAELE DEL MONTE TABOR.

XX PI Locatelli G, Lusso P, Mainati M, Salvatori F, Scarlatti G;

XX DR WPI; 2000-387819/33.

XX Quantitative detection of nucleic acids, based on the use of a
 PT calibrator, suitable primers and probes, and a nucleic acid polymerase
 PT with 5'-3' nuclease activity

XX Example 1; Page 13; 39pp; English.

XX A new method for the quantitative detection of nucleic acids in a
 CC sample which uses a calibrator, suitable primers and probes, and
 CC a nucleic acid polymerase with 5'-3' nuclease activity is described.
 CC The method comprises adding a calibrator nucleotide to the sample,
 CC the calibrator nucleotide having the same sequence as the target with
 CC the exception of one or more regions which in the target hybridize
 CC with a probe labeled with a reporter and quencher, or which hybridize
 CC with the probe and with 2 or more primers. These regions have a
 CC different, randomized nucleotide sequence and a similar Tm. Then
 CC extracting the calibrator and target nucleotides from the sample.
 CC Probes and primers are then added to the extracted sample/calibrator
 CC mixture and PCR is performed. Reaction in the presence of the target
 CC nucleic acid specific probe permits quantitation of the copy number

CC of the extracted target nucleic acid. The reaction in the presence of
 CC the calibrator nucleotide permits quantitation of the calibrator copy
 CC number. The reaction in the presence of both permits calculation of
 CC the total number of target templates and calibrator, allowing
 CC calculation of the percentage of calibrator recovery. The method may
 CC be used for the diagnosis of viral and any other pathogenic agents in
 CC body fluids, and to monitor safety and/or genetic composition of
 CC waters, foods, and plant species used in the alimentary field. The
 CC method is used to quantitatively detect the genomic nucleic acid of
 CC human herpes virus (HHV)-6, HHV-7, HHV-8 and human immunodeficiency
 CC virus (HIV). The method provides an enhanced sensitivity, accuracy,
 CC and precision and a reduced measure viability. For specific probe,
 CC primer and calibrator sequences used in the method, see GENESEQ
 XX records AAZ93935-293965.

SQ Sequence 27 BP; 5 A; 7 C; 7 G; 8 T; 0 other;

Query Match 79.0%; Score 15.8; DB 21; Length 27;

Best Local Similarity 89.5%; Pred. No. 28;

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ctatagtgctaggtatctgt 19

|||||

Db 5 ctatagtgctaggtatctgt 23

RESULT 3

AAZ65560

ID AAZ65560 standard; DNA; 15 BP.

XX AC AAZ65560;

XX DT 12-FEB-2001 (first entry)

XX DE Human focal adhesion kinase antisense sequence #26.

XX KW Human; focal adhesion kinase; FAK; signal transduction; cancer;

XX KW embryonic development disorder; angiogenic disorder; wound healing;

XX KW antisense; phosphorothioate; ss.

XX OS Homo sapiens.

XX PN US6133031-A.

XX PD 17-OCT-2000.

XX PF 19-AUG-1999; 99US-0377310.

XX PR 19-AUG-1999; 99US-0377310.

XX PA (ISIS-) ISIS PHARM INC.

XX PI Monia BP, Gaarde WA;

XX DR WPI; 2001-006141/01.

XX New antisense compounds for inhibiting focal adhesion kinase

PT expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

XX Example 2; Column 25; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX Sequence 15 BP; 4 A; 2 C; 4 G; 5 T; 0 other;

Query Match 75.0%; Score 15; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 agatgctaggtatct 17
 |||||
 Db 1 agatgctaggtatct 15

RESULT 4
 AAC65541
 ID AAC65541 standard; DNA; 20 BP.
 XX
 AC AAC65541;
 XX
 DT 12-FEB-2001 (first entry)
 XX
 DE Human focal adhesion kinase antisense sequence #7.
 XX
 DE Human; focal adhesion kinase; FAK; signal transduction; cancer;
 KW embryonic development disorder; angiogenic disorder; wound healing;
 KW antisense; phosphorothioate; ss.
 XX
 OS Homo sapiens.
 XX
 PN US6133031-A.
 XX
 PD 17-OCT-2000.
 XX
 PF 19-AUG-1999; 99US-0377310.
 XX
 PR 19-AUG-1999; 99US-0377310.
 XX
 XX (ISIS-) ISIS PHARM INC.
 PA
 PI Monia BP, Gaarde WA;
 XX
 DR WPI; 2001-006141/01.
 XX
 XX New antisense compounds for inhibiting focal adhesion kinase
 PT expression, especially useful for inhibiting retinal
 PT neovascularization, or for diagnosing and treating e.g. colon cancer -
 XX
 PS Claim 3; Column 23; 30pp; English.
 XX
 CC The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.
 XX
 SQ Sequence 20 BP; 4 A; 2 C; 5 G; 9 T; 0 other;

Query Match 75.0%; Score 15; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctatagctaggtat 15
 |||||
 Db 6 ctatagctaggtat 20

RESULT 5
 AAQ50213/c
 ID AAQ50213 standard; DNA; 43 BP.
 XX
 AC AAQ50213;
 XX
 DT 06-MAY-1994 (first entry)

Query Match 75.0%; Score 15; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctatagctaggtat 15
 |||||
 Db 6 ctatagctaggtat 20

RESULT 6
 AAQ28274
 ID AAQ28274 standard; cDNA; 38 BP.
 XX
 AC AAQ28274;
 XX
 DT 15-FEB-1993 (first entry)
 XX
 DE PCR primer 2 for a novel type III RTK gene - the KDR gene.
 KW Receptor tyrosine kinase; vascular endothelial cell growth factors;
 KW cancer; tumour; diagnosing; monitoring; ss.

XX
 DE C-fos position 3392-3434 INS mutagenic oligonucleotide.
 XX
 KW HIV-1; p17gag; inhibition; gag; M1; M2; M3; M4; silent mutation; mRNA;
 KW vector; p17; point mutation; p17M1234; Hltat cells; gene replacement;
 KW inhibitory/instability signal; INS; stability; utilisation; vaccine;
 KW interferon; interleukin; fos proto-oncogene protein; growth factor;
 KW env; attenuated; AIDS; AIDS-related disease; latent infection;
 KW gene therapy; Human immunodeficiency virus type 1; p24;
 KW long terminal repeat; LTR; ss.
 XX
 OS Synthetic.
 XX
 PN WO9320212-A.
 XX
 PD 14-OCT-1993.
 XX
 PF 29-MAR-1993; 93WO-US02908.
 XX
 PR 27-MAR-1992; 92US-0858747.
 XX
 XX (USSH) US SEC DEPT HEALTH.
 PA
 PI Felber BK, Pavlakis GN;
 XX
 DR WPI; 1993-336919/42.
 XX
 PT Eliminating inhibitory-instability regions in mRNA to improve
 PT stability and expression - by making multiple point mutations
 PT within A-T rich regions
 XX
 PS Example 3; Page 69; 117pp; English.
 XX
 CC The sequences given in AAQ50202-70 are oligonucleotides which were
 CC used to illustrate the method of the invention for the mutagenesis of
 CC inhibitory/instability signals (INS). Mutation in an INS produced
 CC using oligonucleotides such as these, increases the stability and/or
 CC utilisation of mRNA without changing its protein coding capacity, or
 CC if the sequence is changed, its function is maintained. Other genes
 CC encoding such mRNA molecules include growth factor, interferon,
 CC interleukin, fos proto-oncogene protein and HIV env and gag gene
 CC proteins. Nucleic acid constructs in which INS function has been
 CC impeded, can be used as vaccines, esp. against AIDS and AIDS-related
 CC diseases by preventing HIV from establishing a latent infection, as
 CC is possible using the INS, and thus escaping immune system
 CC surveillance. The constructs may also be used in gene therapy for
 CC gene replacement by homologous recombination with a target gene in
 CC situ. See also AAQ50200-02.
 XX
 SQ Sequence 43 BP; 9 A; 11 C; 5 G; 18 T; 0 other;

Query Match 67.0%; Score 13.4; DB 14; Length 43;
 Best Local Similarity 93.3%; Pred. NO. 5.1e+02;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 gatgctaggtatctg 18
 |||||
 Db 19 GAAGCTAGGTATCTG 5

RESULT 6
 AAQ28274
 ID AAQ28274 standard; cDNA; 38 BP.
 XX
 AC AAQ28274;
 XX
 DT 15-FEB-1993 (first entry)
 XX
 DE PCR primer 2 for a novel type III RTK gene - the KDR gene.
 KW Receptor tyrosine kinase; vascular endothelial cell growth factors;
 KW cancer; tumour; diagnosing; monitoring; ss.

```

XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT modified_base 15
XX FT /*tag= a
XX FT /mod_base= I
XX PN W09214748-A.
XX PD 03-SEP-1992.
XX XX
XX PF 20-FEB-1992; 92WO-US01300.
XX PR 22-FEB-1991; 91US-0657236.
XX XX (AMCY ) AMERICAN CYANAMID CO.
XX PA Carrion ME, Terman BI;
XX PI WPI; 1992-316117/38.
XX DR
XX PT DNA encoding type III receptor tyrosine kinase - useful for
XX PT diagnosing the onset of cancer
XX XX
XX PS Claim 16; Fig 2; 101pp; English.
XX XX
XX CC This PCR primer is designed from a region of the kinase domain 5'
XX CC to the kinase insert domain of a consensus sequence of known type III
XX CC RTK gene catalytic domains. It is used with AAQ28273 to amplify human
XX CC endothelial cDNA producing 251 and 420 bp products. Sequencing of the
XX CC 251bp product revealed a novel sequence containing both primers but with
XX CC little homology to known tyrosine kinases.
XX XX
XX SQ Sequence 38 BP; 5 A; 10 C; 8 G; 10 T; 5 other;

Query Match 65.0%; Score 13; DB 13; Length 38;
Best Local Similarity 68.4%; Pred. No. 8.1e+02;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ctatgatcgttagtctgt 19
Db 19 ctrgcgcagctctcgt 37

RESULT 7
ID AAT97310 standard; DNA; 24 BP.
XX AC AAT97310;
XX DT
XX DT 06-MAR-1998 (first entry)
XX DE
XX DE Human Fas signal sequence generating oligonucleotide Fas1.
XX KW human Fas ligand; hFasL; fusion protein; truncated; T lymphocyte;
XX KW glycosyl-phosphatidylinositol; GPI; treatment; allograft; xenograft;
XX KW Fas signal sequence; ss.
XX OS
XX OS Synthetic.
XX OS Homo sapiens.
XX PN W09718307-A1.
XX XX
XX PD 22-MAY-1997.
XX PF 15-NOV-1996; 96WO-EP05039.
XX PR 16-NOV-1995; 95GB-0023469.
XX XX (SANO ) SANDOZ LTD.
XX PA (SANO ) SANDOZ PAT.

```

```

PA (SANO ) SANDOZ-ERF.
XX XX
XX PI Buehler T;
XX DR
XX DR WPI; 1997-289283/26.
XX PT Human Fas Ligand fused to carboxy-terminal glycopospholipid -
XX PT useful for preventing or treating tissue or organ allograft or
XX PT xenograft rejection
XX XX
XX PS Example 1; Page 11; 26pp; English.
XX XX
XX CC This oligonucleotide Fas1 is used for the generation of a Fas signal
XX CC sequence for post translational modification of a novel human Fas ligand
XX CC (hFasL) fusion protein. This novel protein comprises of hFasL or a
XX CC truncated or functionally equivalent variant that retains the Fas
XX CC receptor binding and apoptosis inducing properties of hFasL. This is
XX CC linked to glycosyl-phosphatidyl inositol (GPI) at its C-terminus. DNA
XX CC encoding hFasL amino acids, a linker sequence and a human Fas signal
XX CC GPI addition signal sequence derived from human CD16. The fusion protein
XX CC can be produced by culturing COS cells transformed by the expression
XX CC vector product. This novel fusion protein can incorporate its lipid tail
XX CC into cell membranes, e.g. endothelial cell membranes, and thus present
XX CC the FasL protein on the cell surface. This can bind to Fas receptor
XX CC present on other cells, particularly T lymphocytes, and thereby induce
XX CC their apoptosis. This is useful for preventing or treating tissue or
XX CC organ allograft or xenograft rejection. This provides a specific
XX CC treatment for activated T lymphocytes, i.e. only for T lymphocytes that
XX CC expresses the Fas antigen, which attack the transplanted tissue or organ.
XX SQ Sequence 24 BP; 3 A; 5 C; 9 G; 7 T; 0 other;

Query Match 64.0%; Score 12.8; DB 18; Length 24;
Best Local Similarity 87.5%; Pred. No. 9.9e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 agatgctaggtatctg 18
Db 8 agatgctggggatctg 23

RESULT 8
ID AAA37684/c
XX AC AAA37684;
XX AC AAA37684;
XX DT 24-OCT-2000 (first entry)
XX DE
XX DE PCR primer for 1-2-rhamnosyl-transferase cDNA sequence.
XX KW 1-2-rhamnosyl-transferase; hesperidin conversion; orange peel; NHDC;
XX KW flavone-7-O-glucosidase-2'-O-rhamnosyl-transferase; flavanoid glycoside;
XX KW sweetener; neohesperidin dihydrochalcone; grapefruit; pomelo;
XX KW citrus fruit; PCR primer; ss.
XX XX
XX OS Citrus sp.
XX XX
XX PN WO200043490-A2.
XX XX
XX PD 27-JUL-2000.
XX XX
XX PF 20-JAN-2000; 2000WO-IL00038.
XX XX
XX PR 22-JAN-1999; 99IL-0128193.
XX XX (VEDA ) VEDA RES & DEV CO LTD.
XX PA (ISRA ) ISRAEL MIN AGRIC.
XX PI Gressel J, Eyal Y, Fluhr R;
XX XX

```

DR WPI; 2000-499220/44.
 XX Polynucleotide encoding a
 PT flavone-7-O-glucosidase-2"-O-rhamnosyl-transferase, useful for
 PT converting hesperidin from orange peels to the sweetener neohesperidin
 PT dihydrochalcone (NHDC) -
 XX
 PS Example; Page 24; 48pp; English.
 XX
 CC This sequence represents a PCR primer for DNA encoding a
 CC 1'-2-rhamnosyl-transferase. The invention relates
 CC to a polynucleotide encoding a polypeptide with
 CC flavone-7-O-glucosidase-2"-O-rhamnosyl-transferase catalytic activity
 CC (such as 1'-2-rhamnosyl-transferase). The enzyme is useful for converting
 CC hesperidin from orange peels to the sweetener neohesperidin
 CC dihydrochalcone (NHDC), and to provide genetically modified plants of the
 CC Citrus genus including an antisense or sense (for co-suppression)
 CC construct, or knockout integrated construct, to provide less bitter
 CC grapefruits, pomeios and other citrus fruits containing flavanoid
 CC glycosides.
 XX
 SQ Sequence 27 BP; 11 A; 6 C; 5 G; 5 T; 0 other;

Query Match 64.0%; Score 12.8; DB 21; Length 27;
 Best Local Similarity 87.5%; Pred. No. 1e+03;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ctatagctaggtatc 16
 || ||||| |||||
 Db 27 CTTGATGCTTGGTATC 12

RESULT 9
 AAZ87799/C
 ID AAZ87799 standard; DNA; 39 BP.
 XX
 AC AAZ87799;
 XX
 DT 12-MAY-2000 (first entry)
 XX
 DE Human tumour suppressor TMPRSS2 cDNA sequencing primer 1B.
 XX
 KW Tumour suppressor gene; TMPRSS2; cancer; human; drug design;
 KW gene therapy; protein therapy; PCR primer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200000605-A1.
 XX
 PD 06-JAN-2000.
 XX
 PF 29-JUN-1999; 99WO-US14622.
 XX
 PR 29-JUN-1998; 98US-0091044.
 XX
 PA (MYRI-) MYRIAD GENETICS INC.
 XX
 PI Wong AKC, Tavtigian SV, Teng DHF;
 XX
 DR WPI; 2000-170914/15.
 XX
 PT Novel tumor suppressor TMPRSS2 used for the diagnosis and prognosis of
 PT human cancer -
 XX
 PS Example 2; Page 53; 89pp; English.
 XX

The invention provides a new tumour suppressor gene, designated TMPRSS2.
 CC The TMPRSS2 polynucleotides and polypeptides can be used in methods for
 CC diagnosing and prognosing predisposition to cancer in humans. The
 CC polypeptides may also be used in assays to screen for compounds with
 CC anti-cancer or therapeutic properties. The polypeptides are also useful
 CC for rational drug design. The TMPRSS2 polynucleotides and polypeptides

CC may be used for gene therapy and protein therapy. Sequences AAZ87797-812
 CC represent PCR primers for sequencing the TMPRSS2 cDNA.

SQ Sequence 39 BP; 9 A; 11 C; 7 G; 12 T; 0 other;

Query Match 64.0%; Score 12.8; DB 21; Length 39;
 Best Local Similarity 87.5%; Pred. No. 1e+03;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 atgctaggtatctgtc 20
 ||| ||||| |||
 Db 30 ATGATAGGTATCCGTC 15

RESULT 10
 AAT00720
 ID AAT00720 standard; cDNA; 21 BP.
 XX
 AC AAT00720;
 XX
 DT 08-MAY-1996 (first entry)
 XX
 DE Multiple tumour suppressor 1 exon.1 beta gene PCR reverse primer.
 XX
 KW Multiple tumour suppressor; MTS1elbeta; cancer; diagnosis; assay;
 KW predisposition; melanoma; leukaemia; lymphoma; prognosis;
 KW pancreas; breast; thyroid; PCR reverse primer; exon 1; ss.
 XX
 OS Synthetic.
 XX
 PN WO9525813-A1.
 XX
 PD 28-SEP-1995.
 XX
 PF 17-MAR-1995; 95WO-US03537.
 XX
 PR 01-JUN-1994; 94US-0251938.
 PR 18-MAR-1994; 94US-0214582.
 PR 18-MAR-1994; 94US-0215086.
 PR 18-MAR-1994; 94US-0215087.
 PR 14-APR-1994; 94US-0227369.
 XX
 PA (MYRI-) MYRIAD GENETICS INC.
 PA (UTAH) UNIV UTAH RES FOUND.
 XX
 PI Cannon-Allbright LA, Kamb A, Skolnick MH;
 XX
 DR WPI; 1995-344626/44.
 XX
 PT Detecting polymorphism associated with cancer pre:disposition - also
 PT DNA, vectors and host cells e.g. for gene or protein replacement
 PT therapy and drug screening
 XX
 PS Example 10; Page 64; 148pp; English.
 XX
 CC An individual can be diagnosed as having a predisposition to cancer
 CC by detecting an alteration in the wild type multiple tumour
 CC suppressor (MTS) gene, using gene probes which hybridise to the
 CC MTS1elbeta gene ORF (amplified using the PCR primers AAR00719-21).
 CC The above assay can also be used in the diagnosis and prognosis of
 CC melanoma, lymphoma, leukaemia and pancreas, breast and thyroid
 CC cancers, etc..

SQ Sequence 21 BP; 5 A; 3 C; 7 G; 6 T; 0 other;

Query Match 63.0%; Score 12.6; DB 16; Length 21;
 Best Local Similarity 78.9%; Pred. No. 1.2e+03;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagctaggtatctgt 19
 ||||| ||| |||||

```

Db      3 ctaggaggcgaattatctgt 21

RESULT 11
AAQ99928
ID      AAQ99928 standard; cDNA; 21 BP.
AC      AAQ99928;
XX
DT      07-MAY-1996 (first entry)
DE      Human MTS1E1-beta PCR amplification primer.
XX
KW      Multiple tumour suppressor; El-alpha; diagnosis; cancer; leukaemia;
KW      astrocytoma; glioblastoma; Hodgkin's lymphoma; melanoma; glioma;
KW      gene therapy; chronic; ss.
XX
OS      Synthetic.
XX
PN      WO9525429-A1.
XX
PD      28-SEP-1995.
XX
PF      17-MAR-1995; 95WO-US03316.
XX
PR      01-JUN-1994; 94US-0251938.
PR      18-MAR-1994; 94US-0214581.
PR      18-MAR-1994; 94US-0214582.
PR      18-MAR-1994; 94US-0215086.
PR      14-APR-1994; 94US-0227369.
PR      18-MAR-1994; 94US-0215086.
PR      18-MAR-1994; 94US-0215087.
XX
PA      (MYRI-) MYRIAD GENETICS INC.
XX
PI      Kamb A;
XX
WPI: 1995-344401/44.
XX
PT      Wild-type multiple tumour suppressor (MTS) gene and mutant sequences
PT      - useful in diagnosis, prognosis and therapy of human cancer, e.g.
PT      melanoma or leukaemia
XX
PS      Example 10; Page 64; 156pp; English.
XX
CC      The cDNA sequences encoding several multiple tumour suppressor (MTS)
CC      polypeptides have been isolated and sequenced, using various
CC      sequencing and amplification primers such as the primer represented
CC      in this sequence. MTS polypeptide-encoding cDNAs and mutants of
CC      these are useful for the diagnosis or prognosis of human cancer.
CC      Germ-line mutations of MTS cDNAs can be used for diagnosing
CC      predisposition to melanoma, leukaemia, astrocytoma, glioblastoma,
CC      lymphoma, glioma, Hodgkin's lymphoma, CLL and cancers of the pancreas,
CC      thyroid, ovary, uterus, testis, kidney, stomach and rectum. The
CC      wild-type gene is useful for gene therapy and MTS polypeptides may also
CC      be used for protein replacement therapy. Also the polypeptides or cells
CC      contg. an altered MTS gene are useful for screening for potential
CC      cancer therapeutics.
XX
SQ      Sequence 21 BP; 5 A; 3 C; 7 G; 6 T; 0 other;

Query Match      63.0%; Score 12.6; DB 16; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 ctagatgctaggtagtctgt 19
        ||||| || | |||||
Db      3 ctaggaggcgaattatctgt 21

RESULT 12
AAT72309
ID      AAT72309 standard; DNA; 21 BP.
AC      AAT72309;
XX
DT      10-SEP-1997 (first entry)
DE      Human multiple tumour suppressor gene 1 Elbeta reverse primer.
XX
KW      Primer; polymerase chain reaction; PCR; amplification; Elbeta;
KW      human; multiple; tumour; suppressor; MTS1; cancer; diagnosis; ss.
XX
OS      Synthetic.
XX
PN      US5624819-A.
XX
PD      29-APR-1997.
XX
PF      18-MAR-1994; 94US-0214582.
XX
PR      07-JUN-1995; 95US-0474177.
PR      18-MAR-1994; 94US-0214582.
PR      18-MAR-1994; 94US-0215086.
PR      18-MAR-1994; 94US-0215087.
PR      14-APR-1994; 94US-0227369.
PR      01-JUN-1994; 94US-0251938.
PR      17-MAR-1995; 95WO-US03537.
XX
PA      (MYRI-) MYRIAD GENETICS INC.
PA      (UTAH ) UNIV UTAH RES FOUND.
XX
PI      Cannon-Albright LA, Kamb A, Skolnick MH;
XX
WPI: 1997-258217/23.
XX
PT      Human mutant multiple tumour suppressor gene sequences - for
PT      production of recombinant mutant polypeptide(s)
XX
PS      Example 10; Columns 87-88; 72pp; English.
XX
CC      The present sequence is primer for the PCR amplification of the
CC      human multiple tumour suppressor gene 1 Elbeta (MTS1Elbeta),
CC      useful in cancer diagnosis.
XX
SQ      Sequence 21 BP; 5 A; 3 C; 7 G; 6 T; 0 other;

Query Match      63.0%; Score 12.6; DB 18; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy      1 ctagatgctaggtagtctgt 19
        ||||| || | |||||
Db      3 ctaggaggcgaattatctgt 21

RESULT 13
AAV53849
ID      AAV53849 standard; DNA; 21 BP.
AC      AAV53849;
XX
DT      04-DEC-1998 (first entry)
DE      Nucleotide sequence of PCR primer 13.
XX
KW      Multiple tumour suppressor; MTS; human; cancer; hybridisation;
KW      somatic mutation; gene therapy; PCR; primer; amplification; ss.
XX
OS      Synthetic.
XX
PN      US5801236-A.
XX
PD      01-SEP-1998.

```

```

XX PF 07-JUN-1995; 95US-0480810.
XX PR 07-JUN-1995; 95US-0480810.
XX PR 18-MAR-1994; 94US-0214582.
XX PR 18-MAR-1994; 94US-0215086.
XX PR 18-MAR-1994; 94US-0215087.
XX PR 14-APR-1994; 94US-0227369.
XX PR 01-JUN-1994; 94US-0251938.
XX PR 01-JUN-1994; 94US-0251938.
XX PR 17-MAR-1995; 95WO-US03316.
XX PA (MYRI-) MYRIAD GENETICS INC.
XX PI Kamb A;
XX DR WPI; 1998-494842/42.
XX PT Nucleic acids based on multiple tumour suppressor, MTS, sequences -
XX PT useful as hybridisation probes, primers and recombinant production
XX PT of MTS in the diagnosis and treatment of cancers related to MTS
XX PT mutation(s)
XX PS Example 10; Column 87-88; 73pp; English.
XX CC This is the nucleotide sequence of a PCR primer used for
XX CC amplification in the method of the invention involving the use of
XX CC the multiple tumour suppressor (MTS) gene, to diagnose and treat
XX CC cancer. The MTS gene is useful in the diagnosis and prognosis of
XX CC human cancer, e.g. by standard nucleic hybridisation techniques, of
XX CC patient samples. The mutated sequences are those that are present
XX CC in somatic mutations of the gene in cancers. The vectors can be
XX CC used for gene therapy strategies to replace function of mutated
XX CC protein in patients. These can also be used to construct protein
XX CC mimetics, also for therapeutic strategies. In addition the
XX CC expression constructs can also be used for recombinant production
XX CC of MTS. Recombinant MTS can be used to screen for drugs to be used
XX CC for cancer therapy, and the protein itself may also be used to
XX CC restore MTS function in a cell.
XX SQ Sequence 21 BP; 5 A; 3 C; 7 G; 6 T; 0 other;

Query Match 63.0%; Score 12.6; DB 19; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagctcgtaggctatctgt 19
    ||||| || | |||||
Db 3 ctaggaggcgaattatctgt 21

RESULT 14
AAV11268
ID AAV11268 standard; DNA; 21 BP.
AC AAV11268;
XX 15-JUL-1998 (first entry)
DE Human MTS1E1-beta PCR primer #2.
XX MTS1E1-beta; multiple tumour suppressor; diagnosis; cancer;
KW germ-line mutation; familial melanoma locus; MLM; predisposition; ss.
XX Synthetic.
OS Homo sapiens.
XX US5739027-A.
XX 14-APR-1998.
XX 07-JUN-1995; 95US-0487033.
XX

Query Match 63.0%; Score 12.6; DB 19; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctatagctcgtaggctatctgt 19
    ||||| || | |||||
Db 3 ctaggaggcgaattatctgt 21

RESULT 15
AAV70613
ID AAV70613 standard; DNA; 21 BP.
AC AAV70613;
XX 03-FEB-1999 (first entry)
DE Reverse PCR primer used to amplify human MTS1E1-beta gene.
XX Human; multiple tumour suppressor 1 gene; MTS1; cancer; PCR primer; ss.
XX Synthetic.
OS Homo sapiens.
XX US5843756-A.
XX 01-DEC-1998.
XX 28-JUL-1995; 95WO-0508735.
XX 28-JUL-1995; 95US-0508735.
XX 07-JUN-1995; 95US-0487033.
XX (MYRI-) MYRIAD GENETICS INC.
XX Jiang P, Kamb A, Stone S;
XX WPI; 1999-044585/04.
XX Mouse multiple tumour suppressor gene segment - useful for primer
XX design

```

XX Example 11; Column 48; 80pp; English.
PS
XX
CC PCR primers AAV70612-14 were used to amplify nucleic acid encoding a
CC human multiple tumour suppressor 1E1-beta (MTE1-beta) protein.
CC Primers designed from the gene can be used to design primers to
CC detect abnormalities i.e. polymorphisms which may predispose
CC towards malignancies such as melanoma, leukaemia, astrocytoma,
CC lymphoma, glioma, as well as tumours of e.g. the breast, thyroid,
CC pancreas, uterus and kidneys.
XX
SQ Sequence 21 BP; 5 A; 3 C; 7 G; 6 T; 0 other;

Query Match 63.0%; Score 12.6; DB 20; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.2e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ctagatgctagggtatctgt 19
 ||||| | | |||||
Db 3 ctgagggcgattatctgt 21

Search completed: October 2, 2001, 16:18:38
Job time: 15482 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:36 ; Search time 876.95 seconds
(without alignments)
14.320 Million cell updates/sec

Title: US-09-757-100B-7

Perfect score: 20

Sequence: 1 ttctcccttcgattattctt 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	22	AAC65539 Human focal adhesi
2	15.2	76.0	36	20	AA04289 Human prostate-spe
3	15	75.0	15	22	AAC65559 Human focal adhesi
4	13.8	69.0	33	19	AAV49869 LM609 grafted anti
5	13.8	69.0	33	22	AAF28225 DNA encoding mutan
6	13.4	67.0	46	19	AAV27760 M13mp18 template s
7	13.2	66.0	29	21	AA04352 Polymorphic fragme
8	12.8	64.0	17	16	AAQ97859 MDV L1 CDNA primer
9	12.8	64.0	16	20	AA01237 PCR primer for PGI
10	12.8	64.0	21	20	AA056167 Human alpha-7 nico
11	12.6	63.0	20	20	AA206154 PCR primer used to

12	12.6	63.0	20	20	AA036892 Human XLIS gene fr
13	12.6	63.0	24	20	AA230696 A. oryzae 40S ribo
14	12.6	63.0	34	17	AA242899 ss circular oligo
15	12.6	63.0	41	19	AAV50562 Brassica sp. polym
16	12.6	63.0	41	19	AAV50563 Brassica sp. polym
17	12.4	62.0	25	21	AAAG68425 Bacteriophage 3A O
18	12.4	62.0	25	21	AAAG68418 Bacteriophage 96 O
19	12.4	62.0	25	21	AAAG68819 Bacteriophage 96 O
20	12.2	61.0	21	19	AA25981 Human polymorphic
21	12.2	61.0	22	18	AA25981 Human tub gene 5'
22	12.2	61.0	22	21	AA25981 PCR primer #2 for
23	12.2	61.0	32	16	AA25405 Human gene signatu
24	12.2	61.0	32	21	AA09244 Primer SIRrev to c
25	12.2	61.0	48	21	AA293280 Oligonucleotide us
26	12.2	61.0	48	21	AA293281 Oligonucleotide us
27	12	60.0	21	22	AA295547 Human gene single
28	12	60.0	26	21	AA22241 Arabidopsis thalia
29	12	60.0	26	21	AA091678 I-SceI gene partia
30	12	60.0	27	21	AA091680 I-SceI gene partia
31	12	60.0	30	14	AA053647 Nucleic acid ligan
32	12	60.0	30	14	AA052035 B-cell mRNA ribozy
33	12	60.0	30	17	AA07685 NGF SELEX clone ha
34	12	60.0	30	18	AAV00775 SELEX ligand hsnf
35	12	60.0	30	19	AAV64599 Human native inter
36	12	60.0	30	19	AAV14556 SELEX identified l
37	12	60.0	30	20	AAV79635 RNA ligand sequenc
38	12	60.0	30	21	AA092735 Nerve growth facto
39	12	60.0	34	17	AA042888 ss circular oligo
40	12	60.0	39	20	AA019124 Tomato prosystemin
41	12	60.0	39	20	AA069995 Elastase DNA nucle
42	12	60.0	40	16	AA089104 Human brx PCR prim
43	12	60.0	40	20	AA024049 PCR primer oligo D
44	12	60.0	43	16	AA01299 DNA encoding unive
45	12	60.0	45	18	AA070280

ALIGNMENTS

RESULT 1
AAC65539
ID AAC65539 standard; DNA; 20 BP.
XX
AC AAC65539;
XX
XX 12-FEB-2001 (first entry)
XX Human focal adhesion kinase antisense sequence #5.
DE Human; focal adhesion kinase; FAK; signal transduction; cancer;
KW embryonic development disorder; angiogenic disorder; wound healing;
KW antisense; phosphorothioate; ss.
XX
OS Homo sapiens.
XX
XX US6133031-A.
XX
PD 17-OCT-2000.
XX
XX 19-AUG-1999; 99US-0377310.
PF
XX 19-AUG-1999; 99US-0377310.
PR
XX (ISIS-) ISIS PHARM INC.
PA
XX Monia BP, Gaarde WA;
PI
XX DNA encoding mutan
XX M13mp18 template s
DR
XX Polymorphic fragme
XX MDV L1 CDNA primer
PT PCR primer for PGI
PT Human alpha-7 nico
PT PCR primer used to

PS Claim 3; Column 23; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense sequences to the human focal adhesion kinase (FAK) protein. This protein is involved in integrin-mediated signal transduction, and is implicated in cancer, particularly colon, breast and oral tumours, embryonic development disorders, angiogenic disorders and wound healing. The antisense sequences, including the one shown here, can be used in the treatment of all of these.

XX Sequence 20 BP; 1 A; 7 C; 1 G; 11 T; 0 other;

SQ

Query Match 100.0%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.68;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
 ||||| ||||| ||||| |||||

Db 1 ttctcccttcggtattctt 20

RESULT 2

AAx04289
 ID AAX04289 standard; DNA; 36 BP.

XX
 AC AAX04289;

XX
 DT 15-APR-1999 (first entry)

XX
 DE Human prostate-specific antigen linker regions of PAP-290.

XX Ricin-like toxin; cancer; viral infection; parasitic infection;
 KW linker; B chain; A chain; protease; fungal infection; malaria;
 KW leucocyte proliferation; cytomegalovirus; herpes; hepatitis;
 KW rhinovirus; laryngeotracheitis; poliomyelitis; varicella zoster;
 KW cystic fibrosis; multiple sclerosis; ds.

XX
 OS Homo sapiens.

OS Synthetic.

XX
 PN W09849311-A2.

XX
 PD 05-NOV-1998.

XX
 PF 30-APR-1998; 98WO-CA00394.

XX
 PR 29-OCT-1997; 97US-0063715.

PR 30-APR-1997; 97US-0045148.

XX
 PA (DNOV-) DE NOVO ENZYME CORP.

XX
 PI Borgford T;

XX
 DR WPI; 1999-009431/01.

XX
 PT New nucleic acid encoding ricin-like toxin with an interchain linker
 PT cleaved by protease - is specific for diseased cells, useful for,
 PT e.g. killing selectively cancer or infected cells

XX
 PS Disclosure; Fig 44B; 352pp; English.

XX
 CC The present invention describes new purified and isolated nucleic acids
 CC (I) encoding: (i) the A and B chains of a ricin-like toxin (II); and
 CC (ii) a heterologous linker, joining the two chains and including a
 CC cleavage recognition site for a disease-specific protease (III). Also
 CC described are: (1) plasmids or baculovirus transfer vectors that contain
 CC of (I); and (2) recombinant protein (IV) consisting of the A and B chains
 CC of (II) joined by the specified linker. (IV), produced by expression of
 CC (I) in host cells, are used to inhibit or kill diseased cells that
 CC produce (III), particularly for treating cancers (e.g. leucocyte
 CC proliferation; cancer of ovary, pancreas, breast or prostate; glioma) or
 CC infections caused by fungi, parasites (e.g. malaria) or viruses (e.g.

CC cytomegalovirus (CMV), herpes, hepatitis, rhinovirus, laryngeotracheitis,
 CC poliomyelitis or varicella zoster), also cystic fibrosis and multiple
 CC sclerosis. Alternatively, (I) is used to express (IV) in vivo. (IV) is
 CC toxic specifically for (III)-expressing cells and does not depend for
 CC specificity on a cell-binding component. When used to treat virus-
 CC infected cells, transcytosis and cytotoxicity of (IV) are increased by
 CC retrograde translocation from endoplasmic reticulum to cytoplasm (which
 CC some viruses exploit to avoid immune detection), so selectivity and
 CC safety are further improved. (IV) are not toxic until chain A is
 CC released and this occurs only in target cells. The present sequence
 CC represents a nucleotide sequence from the present invention.

XX
 SQ Sequence 36 BP; 5 A; 9 C; 3 G; 19 T; 0 other;

Query Match 76.0%; Score 15.2; DB 20; Length 36;
 Best Local Similarity 85.0%; Pred. No. 1.4e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ttctcccttcggtattctt 20
 ||||| ||||| ||||| |||||

Db 14 ttctcccttcggtattctt 33

RESULT 3

AAc65559
 ID AAC65559 standard; DNA; 15 BP.

XX
 AC AAC65559;

XX
 DT 12-FEB-2001 (first entry)

XX
 DE Human focal adhesion kinase antisense sequence #25.

XX
 KW Human; focal adhesion kinase; FAK; signal transduction; cancer;
 KW embryonic development disorder; angiogenic disorder; wound healing;
 KW antisense; phosphorothioate; ss.

XX
 OS Homo sapiens.

XX
 PN US6133031-A.

XX
 PD 17-OCT-2000.

XX
 PF 19-AUG-1999; 99US-0377310.

XX
 PR 19-AUG-1999; 99US-0377310.

XX
 PA (ISIS-) ISIS PHARM INC.

XX
 PI Monia BP, Gaarde WA;

XX
 DR WPI; 2001-006141/01.

XX
 PT New antisense compounds for inhibiting focal adhesion kinase
 PT expression, especially useful for inhibiting retinal
 PT neovascularization, or for diagnosing and treating e.g. colon cancer -

XX
 PS Example 2; Column 25; 30pp; English.

XX
 CC The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX
 SQ Sequence 15 BP; 1 A; 6 C; 1 G; 7 T; 0 other;

Query Match 75.0%; Score 15; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 ctccctccgtatt 17
|||||
Db 1 ctccctccgtatt 15

RESULT 4

AAV49869
ID AAV49869 standard; DNA; 33 BP.

XX AAV49869;

XX 02-NOV-1998 (first entry)

XX LM609 grafted antibody V-L region CDR2 DNA fragment #2.

XX Vitaxin; antibody; variable region; heavy chain; light chain; integrin;
KW LM609; inhibitor; integrin-mediated signal transduction; treatment;
KW diagnosis; angiogenesis; restenosis; inflammation; diabetic retinopathy;
KW neovascular glaucoma; cancer; psoriasis; rheumatoid arthritis; murine;
KW macular degeneration; osteoporosis; primer; V-L region; CDR;
KW complementarity determining region; ss.

OS Mus sp.

XX WO9833919-A2.

XX 06-AUG-1998.

XX 30-JAN-1998; 98WO-US01826.

XX 30-JAN-1997; 97US-0791391.

XX (IXSY-) IXSYS INC.

XX Glaser SM, Huse WD;

XX WPI; 1998-437472/37.

XX P-PSDB; AAW76032.

XX Humanised antibody, Vitaxin, that binds selectively to alphaVbeta3
PT integrin - and related grafted antibodies based on murine monoclonal
PT LM609, also related nucleic acid, used to treat, prevent or diagnose
PT angiogenesis or restenosis

XX Disclosure; Page 42; 129pp; English.

XX AAV49844-V49877 are nucleotide fragments of the grafted monoclonal
CC antibody LM609 heavy and light chain variable region. LM609 and the
CC antibody vitaxin bind selectively to integrin alphavbeta3 and can be used
CC to inhibit binding of alphavbeta3 to a ligand and thus block
CC integrin-mediated signal transduction. This is useful in the treatment,
CC prevention and diagnosis of alphavbeta3-mediated disease, specifically
CC angiogenesis and restenosis (but also e.g. (non-)immune inflammation,
CC diabetic retinopathy, neovascular glaucoma, cancer, psoriasis,
CC rheumatoid arthritis, macular degeneration, osteoporosis etc.). The
CC antibodies contain non-murine framework regions so are suitable for use
CC in humans. Enhanced types of LM609 have affinity more than 90 times
CC greater than that of parent the parent antibody.

XX Sequence 33 BP; 4 A; 13 C; 2 G; 14 T; 0 other;

Query Match 69.08; Score 13.8; DB 19; Length 33;
Best Local Similarity 88.2%; Pred. No. 6.5e+02;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 4 tcctctccgtattctt 20
|||||
Db 3 tctcatccgtattctt 19

RESULT 5

AAF28225
ID AAF28225 standard; DNA; 33 BP.

XX AAF28225;

XX 03-APR-2001 (first entry)

XX DNA encoding mutant VL CDR2 peptide.

XX LM609; grafted antibody; alphaVbeta3 integrin; angiogenesis;
KW inflammatory; cancer; retina; restenosis; osteoporosis; ss.

XX Unidentified.

XX WO200078815-A1.

XX 28-DEC-2000.

XX 23-JUN-2000; 2000WO-US17454.

XX 24-JUN-1999; 99US-0339922.

XX (MOLE-) APPLIED MOLECULAR EVOLUTION.

XX Huse WD, Wu H;

XX WPI; 2001-050110/06.

XX Enhanced LM609 grafted antibodies exhibiting selective binding affinity
PT to alpha(V)beta(3) integrin, useful in the diagnosis and treatment of
PT angiogenesis, inflammatory diseases e.g. psoriasis, cancers and
PT osteoporosis

XX Disclosure; Page 42; 132pp; English.

XX The present invention relates to enhanced LM609 grafted antibodies
CC exhibiting selective binding affinity to alphaVbeta3 integrin or
CC their functional fragments. The antibodies or their functional
CC fragments can be used in the diagnosis and treatment of
CC alphaVbeta3-mediated diseases such as angiogenesis, inflammatory
CC diseases (such as psoriasis and chronic articular rheumatism),
CC disorders associated with inappropriate or inopportune invasion of
CC vessels (such as diabetic retinopathy, neovascular glaucoma and
CC cancer disorders such as tumours and Kaposi's sarcoma), retinal
CC diseases (such as macular degeneration), restenosis and
CC osteoporosis.

XX Sequence 33 BP; 4 A; 13 C; 2 G; 14 T; 0 other;

Query Match 69.08; Score 13.8; DB 22; Length 33;

Best Local Similarity 88.2%; Pred. No. 6.5e+02; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 2;

Oy 4 tcctctccgtattctt 20
|||||
Db 3 tctcatccgtattctt 19

RESULT 6

AAV27760
ID AAV27760 standard; DNA; 46 BP.

XX AAV27760;

XX 24-SEP-1998 (first entry)

XX M13mp18 template sequence.

XX ss; M13mp18; sequencing; target specificity; PCR; amplification;
KW differential primer extension.

XX

OS Bacteriophage M13.
 XX
 FH Key Location/Qualifiers
 FT misc_feature 20..21
 FT /*tag= a
 FT /note= "Unspecified amount of nucleotides not given
 FT in specification"
 XX
 XX
 PN W09814608-A1.
 XX
 XX
 PD 09-APR-1998.
 XX
 XX 30-SEP-1997; 97WO-US17305.
 XX
 XX 02-OCT-1996; 96IL-0119342.
 XX
 XX (USAT) US DEPT ENERGY.
 PA
 XX Mugasimangalam RC, Ulanovsky LE;
 XX WPI; 1998-286423/25.
 DR
 XX Improving priming specificity of oligo:nucleotide primers - by using
 PT differential extension of the primer with nucleotide subsets and
 PT length discrimination against shorter extensions
 XX
 XX Example 1; Fig 1; 46pp; English.
 XX
 CC The target nucleotide sequence was used to demonstrate a method for
 CC improving the priming specificity of a non-unique oligonucleotide primer.
 CC Preferably the oligonucleotide primer is modified by a radioactive
 CC element, a fluorescent moiety or a biotin moiety. The methods are used
 CC particularly for the sequencing of target DNA sequences. The methods
 CC improve the priming specificity of a non-unique oligonucleotide primer by
 CC using differential extension of the primer with nucleotide subsets and
 CC length discrimination against shorter extensions.
 XX
 XX Sequence 46 BP; 6 A; 19 C; 2 G; 19 T; 0 other;

Query Match 67.0%; Score 13.4; DB 19; Length 46;
 Best Local Similarity 93.3%; Pred. No. 1.1e+03;
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 tccctccggtattc 18
 ||||| |||||
 Db 5 tccctccggtattc 19

RESULT 7
 AAA04352/c
 ID AAA04352 standard; DNA; 29 BP.
 XX
 AC AAA04352;

DT 22-MAY-2000 (first entry)

DE Polymorphic fragment of hypertension associated gene HAPT.

XX Polymorphism; hypertension; agammaglobulinemia; diabetes insipidus;
 KW Lesch-Nyhan syndrome; muscular dystrophy; Wiskott-Aldrich syndrome;
 KW Fabry's disease; familial hypercholesterolemia; hereditary spherocytosis;
 KW polycystic kidney disease; von Willebrand's disease; forensis; human;
 KW tuberosus sclerosis; hereditary hemorrhagica telangiectasia;
 KW familial colonic polyposis; osteogenesis imperfecta; porphyria;
 KW Ehlers-Danlos syndrome; ss.

XX Homo sapiens.

XX Ep955382-A2.

XX 10-NOV-1999.

XX

PF 07-MAY-1999; 99EP-0250150.
 XX
 PR 07-MAY-1998; 98US-0084641.
 PR 03-MAY-1999; 99US-0304232.
 XX
 PA (AFFY-) AFFYMETRIX INC.
 PA (UYCA-) UNIV CASE WESTERN RESERVE.
 XX
 XX Fan JB, Chakravarti A, Haluska MK;
 PI WPI; 2000-107928/10.
 DR
 XX Novel nucleic acids containing polymorphisms used in the diagnosis of
 PT hypertension -
 PT
 XX Claim 1; Page 33; 53pp; English.
 PS
 XX The invention provides polymorphic fragments of genes associated with
 CC hypertension. The nucleic acids including the polymorphic sites can be
 CC used as probes or primers for expressing variant proteins. Detection of
 CC the polymorphisms is useful in designing prophylactic and therapeutic
 CC regimes customized to underlying abnormalities. The polymorphisms can be
 CC used for association studies for hypertension, and in hypertension can be
 CC diagnostic assays. Where the polymorphisms have strong correlation with
 CC hypertension, within a gene, they are likely to have a causative role in
 CC hypertension. This information can be used to find the precise role of a
 CC polymorphism in the disease, and this can be used to identify potential
 CC drugs which combat the disease. The polymorphisms can be tested for
 CC association with other diseases e.g. agammaglobulinemia, diabetes
 CC insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich
 CC syndrome, Fabry's disease, familial hypercholesterolemia, polycystic
 CC kidney disease, hereditary spherocytosis, von Willebrand's disease,
 CC tuberosus sclerosis, hereditary hemorrhagica telangiectasia, familial
 CC colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, and
 CC acute intermittent porphyria. The polymorphic forms can also be used in
 CC forensics to identify individuals.
 XX
 XX Sequence 29 BP; 11 A; 2 C; 12 G; 3 T; 1 other;

Query Match 66.0%; Score 13.2; DB 21; Length 29;
 Best Local Similarity 75.0%; Pred. No. 1.2e+03;
 Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ttctccctccggtattctt 20
 ||||| :||| |||||
 Db 22 TTCTCCAKTTCGTATTCT 3

RESULT 8
 AAQ97859/c
 ID AAQ97859 standard; cDNA; 17 BP.
 XX
 AC AAQ97859;

DT 06-DEC-1995 (first entry)

DE MDV L1 cDNA primer 1.

XX Marek's disease virus; MDV; vaccine; antisense; therapy; latency;
 KW tumor; primer; polymerase chain reaction; PCR; ss.

XX Synthetic.

XX WO9518860-A.

XX 13-JUL-1995.

XX 05-JAN-1995; 95WO-US00052.

XX 11-JAN-1994; 94US-0180051.

XX (CORR) CORNELL RES FOUND INC.

XX O'Connell PH, Ohashi K, Schat KA;
XX WPI; 1995-255063/33.
XX Marek's disease virus protein and its nucleotide sequence - used to
PT protect chickens against Marek's disease and inhibit the establishment
PT of latency and tumour cell development
XX
XX Example 1; Page 11; 42pp; English.
XX
XX The primers given in AAQ97859-60 were used for the PCR amplification
CC of the upstream region of MDV L1 cDNA. The amplified fragment,
CC Q2L1, was cloned into pBluescript KS+ for sequencing.
XX
XX Sequence 17 BP; 9 A; 2 C; 5 G; 1 T; 0 other;
SQ

Query Match 64.0%; Score 12.8; DB 16; Length 17;
Best Local Similarity 87.5%; Pred. NO. 1.8e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 cccctccgttattctt 20
Db 17 CCCCTCCCTTTATTGTT 2

RESULT 9
AAZ01237/C
ID AAZ01237 standard; DNA; 18 BP.
XX
XX AAZ01237;
XX
XX 27-SEP-1999 (first entry)
XX
XX PCR primer for PGI biallelic markers 4-54-283 and 4-55-95.
XX
XX PGI gene; biallelic marker; PCR primer; PGI-related biallelic marker;
KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen;
KW PSA; human; ss.
XX
XX Synthetic.
OS
OS Homo sapiens.
XX
XX WO9932644-A2.
XX
XX 01-JUL-1999.
XX
XX 22-DEC-1998; 98WO-IB021133.
XX
XX 09-SEP-1998; 98US-0099658.
XX
XX 22-DEC-1997; 97US-0996306.
XX
XX (GEST) GENSET.
XX
XX Blumenfeld M, Bougueleret L, Chumakov I, Cohen D;
XX
XX WPI; 1999-405178/34.
XX
XX Use of a prostate cancer associated gene and biallelic markers
PT derived from it
XX
XX Claim 4; Page 354; 385pp; English.
XX
XX The invention relates to a mammalian PGI gene and protein, and a set of
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are
CC used in a hybridisation assay, a sequencing assay, or in an
CC allele-specific amplification assay for determining the identity of a
CC nucleotide at a PGI-related biallelic marker. The methods can be used to
CC detect and to assess the risk of developing cancer or prostate cancer.
CC Early-stage diagnosis of prostate cancer relies on prostate specific
CC antigen (PSA) dosage. However, the effectiveness of this is limited due
CC to its inability to discriminate between malignant and non-malignant

CC affections of the organ. A need exists for both a reliable diagnostic
CC procedure which would enable early-stage diagnosis, and for preventative
CC and curative treatments of the disease. The PGI gene can be used for
CC detection of prostate cancer, and the risk of developing it in the
XX future, and can also be used to determine therapies for the disease.
SQ Sequence 18 BP; 6 A; 1 C; 8 G; 3 T; 0 other;
Qy 3 ctcctccgttattctt 18
Db 16 CTCCTCCCTCAGTTATTC 1

Query Match 64.0%; Score 12.8; DB 20; Length 18;
Best Local Similarity 87.5%; Pred. NO. 1.8e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 3 ctcctccgttattctt 18
Db 16 CTCCTCCCTCAGTTATTC 1

RESULT 10
AAZ56167/C
ID AAZ56167 standard; DNA; 21 BP.
XX
XX AAZ56167;
XX
XX 15-JUL-1999 (first entry)
XX
XX Human alpha-7 nicotinic receptor PCR primer SEQ ID NO:14.
XX
XX Human; alpha-7 nicotinic receptor; neuronal; hybridisation; probe;
KW alpha-7 neuronal nicotinic acetylcholine receptor; schizophrenia;
KW small cell lung carcinoma; breast cancer; nicotine-dependent illness;
KW epilepsy; juvenile myoclonic epilepsy; Prader-Willi syndrome;
KW Angelman's syndrome; PCR primer; ss.
XX
XX Synthetic.
OS
OS Homo sapiens.
XX
XX WO9920757-A2.
XX
XX 29-APR-1999.
XX
XX 15-OCT-1998; 98WO-US21762.
XX
XX 23-OCT-1997; 97US-0956518.
XX
XX (FREE/) FREEDMAN R.
XX (LEON/) LEONARD S.
XX
XX Freedman R, Leonard S;
XX
XX WPI; 1999-288306/24.
XX
XX Human alpha-7 neuronal nicotinic acetylcholine receptor and related
PT polynucleotides
XX
XX Claim 15; Page 64; 104pp; English.
XX
XX The present invention describes an isolated nucleotide sequence (I)
CC encoding at least a portion of the human alpha-7 neuronal nicotinic
CC acetylcholine receptor (alpha7-hnAChR). Also described are: (1) a
CC peptide encoded by (I); (2) a vector comprising (I); (3) a host cell
CC transformed with a vector of (2); (4) a polynucleotide comprising at
CC least 15 nucleotides which hybridises under stringent conditions to at
CC least a portion of (I); (5) a method for detection of a polynucleotide
CC encoding alpha 7-hnAChR in a biological sample; and (6) a method for
CC amplification of nucleic acid from a sample suspected of containing
CC nucleic acid encoding alpha 7-hnAChR. The primers and probes from the
CC present invention can be used on brain tissue and blood samples of
CC humans suspected of suffering from schizophrenia, small cell lung
CC carcinoma, breast cancer and nicotine-dependent illness. This is
CC particularly useful for diagnosis of schizophrenia. Other illnesses
CC that can be studied/diagnosed are epilepsy (e.g. juvenile myoclonic
CC epilepsy) and Prader-Willi and Angelman's syndromes.

XX SQ Sequence 21 BP; 9 A; 3 C; 8 G; 1 T; 0 other;

Query Match 64.0%; Score 12.8; DB 20; Length 21;
 Best Local Similarity 87.5%; Pred. No. 1.9e+03;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 tctcccttcggtatt 17
 ||||| ||||| ||
 Db 18 TCCTCCCTGGCTTCTT 3

RESULT 11
 AAZ06154
 ID AAZ06154 standard; DNA; 20 BP.
 XX AC AAZ06154;
 XX DT 07-OCT-1999 (first entry)
 XX PCR primer used to amplify an ORF of Chlamydia trachomatis.
 XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
 KW paratrachoma; inclusion conjunctivitis; genital disease; peritrophic;
 KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis; PCR primer;
 KW bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
 XX Synthetic.
 OS Chlamydia trachomatis.
 XX WO9928475-A2.
 XX 10-JUN-1999.
 XX 27-NOV-1998; 98WO-IB01939.
 XX 04-NOV-1998; 98US-0107077.
 XX 28-NOV-1997; 97FR-0015041.
 XX 17-DEC-1997; 97FR-0016034.
 XX (GEST) GENSET.
 XX Griffais R;
 XX WPI; 1999-371125/31.
 XX Genome sequence of Chlamydia trachomatis
 XX Disclosure; Page 1829; 1755pp; English.
 XX PCR primers AAZ01426-Z06209 were used to amplify open reading frames
 CC (ORFs) of the genome of Chlamydia trachomatis (see AAZ01425). These ORFs
 CC encode polypeptides (see AAY36754-Y37949) which can be used as vaccines
 CC against Chlamydia trachomatis. Antisense and ribozyme sequences
 CC can also be used to control growth of the microorganism. Chlamydia
 CC trachomatis is responsible for a large number of diseases, e.g. eye
 CC diseases such as conventional trachoma, nonendemic trachoma,
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
 CC peritrophic, bartholinitis; pneumopathy in breast feeding infants;
 CC and venereal lymphogranulomatosis. The polypeptides of the
 CC invention may be of use in treating these diseases.
 XX Sequence 20 BP; 2 A; 8 C; 2 G; 8 T; 0 other;

Query Match 63.0%; Score 12.6; DB 20; Length 20;
 Best Local Similarity 78.9%; Pred. No. 2.3e+03;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 tctcccttcggtattctt 20
 | |||| | ||||| | |

Db 2 ttcccatcgcgtatccct 20

RESULT 12
 AAX36892
 ID AAX36892 standard; DNA; 20 BP.
 XX AC AAX36892;
 XX 14-JUL-1999 (first entry)
 XX Human XLIS gene fragment PCR primer 6 F.
 XX XLIS gene; human; detection; diagnosis; prenatal diagnosis; therapy;
 KW lissencephaly; LIS; agyria-pachygyria; subcortical laminar heterotopia;
 KW SCLH; cortical dysgenesis; cryptogenic epilepsy; neurological disorder;
 KW neurodegenerative disease; Alzheimer's disease; X-linked disorder;
 KW genetic counselling; PCR primer; ss.
 XX Synthetic.
 OS Homo sapiens.
 XX EP918091-A1.
 XX 26-MAY-1999.
 XX 21-NOV-1997; 97EP-0402811.
 XX 21-NOV-1997; 97EP-0402811.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX Chelly J, Des Portes V, Kahn A, Pinard J;
 XX WPI; 1999-290318/25.
 XX New gene and its gene product expressed in the brain, useful for
 PT diagnosing and treating disorders such as lissencephaly and
 PT subcortical laminar heterotopia
 XX Claim 9; Page 50; 71pp; English.
 XX This sequence is a primer for the human XLIS gene of the invention.
 CC The XLIS fragments may be used to detect abnormalities in the expression
 CC of the XLIS gene transcripts or to compare their sequence with that of
 CC the XLIS transcripts from patients for in vitro especially prenatal
 CC diagnosis of lissencephaly (LIS) (or agyria-pachygyria), subcortical
 CC laminar heterotopia (SCLH), cortical dysgenesis, cryptogenic epilepsies
 CC or neurodegenerative diseases such as Alzheimer's disease. These
 CC disorders mainly affect females as the XLIS gene is X-linked. The XLIS
 CC fragments may also be used to administer to patients to prevent or treat
 CC the above disorders and may be used as a tool in genetic counselling.
 CC Oligonucleotides which bind to the fragments may be used to amplify the
 CC XLIS gene from a sample for comparison to normal samples in the in vitro
 CC diagnosis regime. This may also be performed by amplifying XLIS cDNA from
 CC the mRNA in the sample. Antibodies to XLIS may be used to detect XLIS in
 CC a biological sample or can be administered to patients to prevent or
 CC treat the above disorders. They may also be used to purify XLIS from a
 CC biological sample. XLIS may also be administered to patients to prevent
 CC or treat the above neurological disorders. In addition XLIS may be used
 CC as a marker of neuronal cells at an early stage of development; its
 CC discovery increases understanding of both the neuronal movement which
 CC leads to development of the cortical region of the brain and of the
 CC pathogenesis of the group of neuronal disorders mentioned above.
 XX Sequence 20 BP; 1 A; 7 C; 0 G; 12 T; 0 other;

Query Match 63.0%; Score 12.6; DB 20; Length 20;
 Best Local Similarity 78.9%; Pred. No. 2.3e+03;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattctt 19

Db 2 ttatcccttcctttcttct 20
 || ||||| || |||

RESULT 13

AAZ30696/c
 ID AAZ30696 standard; DNA; 24 BP.

XX AAZ30696;
 AC

XX 15-FEB-2000 (first entry)
 DT XX

DE A. oryzae 40S ribosome protein S28 gene promoter primer.
 XX

XX Promoter; 40S ribosomal protein S28; genetic engineering; amplification;
 KW heterologous protein; gene expression; PCR; primer; ss.
 KW

XX Synthetic.
 OS

XX Aspergillus oryzae.
 OS

XX JP11276170-A.
 PN

XX 12-OCT-1999.
 PD

XX 31-MAR-1998; 98JP-0105712.
 PF

XX 31-MAR-1998; 98JP-0105712.
 PR

XX (AMANO) AMANO PHARM KK.
 XX

PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.
 PA

XX WPI; 1999-626935/54.
 DR

XX A new promoter derived from an Aspergillus genus microbe - useful for
 PT producing exotic proteins
 PT

PS Example 7; Page 5; 11pp; Japanese.
 PS

XX This primer was used to PCR amplify the promoter sequence from the
 CC 40S ribosomal protein S28 gene (AAZ30685) from Aspergillus oryzae. The
 CC invention relates to novel gene promoters (AAZ30680-230685) isolated
 CC from Aspergillus oryzae which can be used in genetic engineering to
 CC express heterologous proteins in Aspergillus.
 CC

XX Sequence 24 BP; 9 A; 6 C; 8 G; 1 T; 0 other;
 SQ

Query Match 63.0%; Score 12.6; DB 20; Length 24;

Best Local Similarity 78.9%; Pred. No. 2.4e+03;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 tctcccttcctgtattctt 20
 ||||| ||| |||

Db 19 TCCTCCCTTCGGTGTGCTT 1
 ||||| ||| |||

RESULT 14

AAZ42899

ID AAT42899 standard; DNA; 34 BP.

XX AAT42899;
 AC

XX 10-JUN-1997 (first entry)
 DT XX

DE ss circular oligo targetted to -385 bcr breakpoint.
 XX

XX single stranded; circular; target sequence; parallel; detection;
 KW binding domain; anti-parallel; loop domain; complementarity; ss;
 KW synthesis; regulation; drug delivery; biosynthesis; tumour cell.
 XX

XX Synthetic.
 OS

XX Key
 FH

Location/Qualifiers

FT misc_feature 1
 FT /*tag= a
 FT /note= "forms 5' > 3' bond with C34"
 FT 1..12
 FT /*tag= b
 FT /note= "forms parallel binding domain"
 FT 18..29
 FT /*tag= c
 FT /note= "forms anti-parallel binding domain"
 FT 34
 FT /*tag= d
 FT /note= "forms 3' > 5' bond with T1"
 FT XX

XX WO9630384-A1.
 PN

XX 03-OCT-1996.
 PD

XX 21-MAR-1996; 96WO-US03757.
 XX

XX 30-MAR-1995; 95US-0413813.
 XX

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.
 PA

XX KOOL ET;
 XX

XX WPI; 1996-455262/45.
 DR

XX Single stranded circular oligo:nucleotide comprising parallel and or
 PT anti-parallel binding domain - used to regulate biosynthesis of
 PT DNA, RNA or protein in targetted mammalian tumour cell in vivo
 XX

XX Example 11; Page 135; 195pp; English.
 PS

XX The sequences given in AAT42898-901 are single stranded (ss) circular
 CC oligonucleotides and their targets, which are used in the inhibition
 CC of the proliferation of myeloid leukaemia cells. These oligos are
 CC specifically targetted to a region in the bcr3/abl2 gene 385
 CC nucleotides 5' to the bcr/abl junction, abd towards the bcr2/abl2
 CC junction. These ss circular oligonucleotides comprise a parallel
 CC binding (P) domain, and/or an anti-parallel binding (AP) domain, and
 CC at least 1 loop domain. The P and AP domains have sufficient
 CC complementarity to bind detectably to 1 strand of a defined nucleic
 CC acid target. The P domain is capable of binding in a parallel manner
 CC to the target. The AP domain is capable of binding in an anti-parallel
 CC manner to the target and the ends of the P and AP domains are separated
 CC by the loop domains. The ss circular oligonucleotides can be used to
 CC regulate the synthesis of DNA, RNA or protein (pref. by DNA replication,
 CC DNA reverse transcription, RNA splicing, RNA polyadenylation, RNA
 CC translocation or protein translocation) by binding a target sequence
 CC in the template. They can also be used to deliver a drug to a specific
 CC cell type by administering a drug covalently bound to them (i.e. to
 CC regulate the biosynthesis of DNA, RNA or protein in a targetted
 CC mammalian tumour cell in vivo, without substantially altering the
 CC biosynthesis of the DNA). They can also be used to detect a target
 CC nucleic acid by detecting an oligonucleotide-target complex. The
 CC circular oligonucleotide can bind both single and double stranded
 CC target nucleic acids, and has enhanced stability, compared to
 CC linear forms.
 XX

SQ Sequence 34 BP; 1 A; 18 C; 1 G; 14 T; 0 other;

Query Match 63.0%;

Best Local Similarity 78.9%; Score 12.6; DB 17; Length 34;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttctcccttcctgtattctt 19
 ||||| ||| |||

Db 7 ttctcccttcctgtattctt 25
 ||||| ||| |||

RESULT 15

AAV50562

ID AAV50562 standard; DNA; 41 BP.
XX AC AAV50562;
XX DT 21-DEC-1998 (first entry)
XX DE Brassica sp. polymorphic marker N2/10B8/N3-2A DNA.
XX KW Polymorphic marker; allele-specific; primer; probe; amplification;
KW hybridisation; plant; hybrid certification; genetic contribution;
XX KW progeny; back-cross; hybrid; ancestry; ss.
XX OS Brassica sp.
XX FH Key Location/Qualifiers
FT variation 21
FT /*tag= a
FT /replace= "c"
FT /note= "polymorphism"
XX PN W09824796-A1.
XX PD 11-JUN-1998.
XX PF 01-DEC-1997; 97WO-US21782.
XX PR 07-MAR-1997; 97US-0813507.
XX PR 02-DEC-1996; 96US-0032069.
XX PA (AFFY-) AFFYMETRIX INC.
XX PI Landry BS, Lemieux B, Murigneux A, Sapolsky RJ;
XX WPI; 1998-333252/29.
XX PT Brassica species allele-specific oligonucleotide probes and primers
XX PT - useful for plant breeding
XX PS Claim 1; Page 32; 65pp; English.
XX CC This DNA sequence is a region of a Brassica napus or Brassica oleracea
CC genome which contains a polymorphic marker. This sequence can be used
CC in the construction of allele-specific primers and probes for
CC amplification or hybridisation, e.g. to determine common or disparate
CC ancestry between 2 or more plants, to monitor the genetic contribution
CC of an ancestral plant, to trace the progeny of proprietary plants, in
CC certification of a hybrid plant or to identify the progeny of a
CC back-crossed plant with an ancestral plant.
XX SQ Sequence 41 BP; 8 A; 11 C; 3 G; 19 T; 0 other;

Query Match 63.0%; Score 12.6; DB 19; Length 41;
Best Local Similarity 78.9%; Pred. No. 2.5e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttctcccttcggtattct 19
Db 12 ttctccagttcttattct 30

Search completed: October 2, 2001, 16:18:37
Job time: 15481 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:45 ; Search time 417.38 Seconds
(without alignments)
9.071 Million cell updates/sec

Title: US-09-757-100B-6

Perfect score: 20

Sequence: 1 gaaactgcagaagcactga 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	3	US-09-377-310-6
2	15	75.0	15	3	US-09-377-310-26
C 3	13.8	69.0	25	1	US-08-288-405A-3
C 4	13.6	68.0	42	4	US-08-935-312-15
C 5	13.4	67.0	43	1	US-08-700-186-12
C 6	13.4	67.0	43	2	US-08-914-981-12
C 7	13.4	67.0	43	3	US-08-116-115-12
8	13.2	66.0	27	1	US-07-825-959-13
9	13.2	66.0	27	1	US-08-131-324-13
10	12.8	64.0	22	1	US-08-377-495-1
11	12.8	64.0	22	2	US-08-748-415-1
12	12.8	64.0	22	3	US-08-747-654-1
13	12.8	64.0	22	4	US-09-235-353-1
14	12.8	64.0	35	4	US-09-232-477-1
15	12.6	63.0	21	1	US-08-105-483-148
16	12.6	63.0	21	1	US-08-709-209-148
17	12.6	63.0	21	1	US-08-303-275-36
18	12.6	63.0	21	1	US-08-458-101-148
19	12.6	63.0	30	2	US-08-442-010-7
20	12.4	62.0	24	3	US-08-487-799-57
21	12.4	62.0	25	1	US-08-488-702-6
22	12.4	62.0	29	4	US-08-972-927-13
C 23	12.4	62.0	39	1	US-07-780-800A-17
C 24	12.4	62.0	39	1	US-08-327-516-4
C 25	12.4	62.0	39	6	5464743-5
C 26	12.4	62.0	39	6	5519127-16
27	12.2	61.0	18	3	US-08-738-381-44

28	12.2	61.0	24	4	US-09-433-694-5
C 29	12.2	61.0	27	1	US-08-180-209B-49
C 30	12.2	61.0	27	5	PCT-US94-02629-49
C 31	12.2	61.0	30	2	US-08-280-546-6
C 32	12.2	61.0	31	2	US-09-036-582-38
C 33	12.2	61.0	36	4	US-09-416-756A-14
C 34	12.2	61.0	42	2	US-08-343-443B-47
C 35	12.2	61.0	50	1	US-08-171-389-396
C 36	12.2	61.0	50	1	US-08-123-936-396
C 37	12.2	61.0	50	2	US-08-475-228A-396
C 38	12.2	61.0	50	3	US-08-482-080A-396
C 39	12.2	61.0	50	5	PCT-US93-12388-396
C 40	12	60.0	23	2	US-08-859-998-118
C 41	12	60.0	26	4	US-09-493-352A-2
C 42	12	60.0	30	1	US-07-979-966A-15
C 43	12	60.0	32	2	US-08-472-659-11
C 44	12	60.0	32	2	US-08-474-661-11
45	12	60.0	32	2	US-08-611-977-11

ALIGNMENTS

RESULT 1
US-09-377-310-6
; Sequence 6, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-6

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gaaactgcagaagcactga 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1 gaaactgcagaagcactga 20

RESULT 2
US-09-377-310-26
; Sequence 26, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence

;
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-26

Query Match 75.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 aactgcagaagcgcac 17
Db 1 aactgcagaagcgcac 15

RESULT 3
US-08-288-405A-3/c
; Sequence 3, Application US/08288405A
; Patent No. 5559009
; GENERAL INFORMATION:
; APPLICANT: Chandy, Kaniyanthara G.
; APPLICANT: Kalman, Katalin
; APPLICANT: Chandy, Grisha
; APPLICANT: Gutman, George A.
; TITLE OF INVENTION: A No. 5559009el Voltage-Gated Potassium Channel
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert,
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,405A
; FILING DATE: 10-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,431
; FILING DATE: 04-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Dregler, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-59844-1/WH/D
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: both
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14..25
US-08-288-405A-3

Query Match 69.0%; Score 13.8; DB 1; Length 25;
Best Local Similarity 88.2%; Pred. No. 3e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gaaactgcagaagcgcac 17
Db 17 GGAAGTGCAGAAGGCAC 1

RESULT 4
US-08-935-312-15/c
; Sequence 15, Application US/08935312
; Patent No. 6207455
; GENERAL INFORMATION:
; APPLICANT: CHANG, Lung-Ji
; TITLE OF INVENTION: LENTIVIRAL VECTORS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,312
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: CHANG-112
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
US-08-935-312-15

Query Match 68.0%; Score 13.6; DB 4; Length 42;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gaaactgcagaagcgcactga 20
Db 42 GAAAGAGCAGAAGACAGTGA 23

RESULT 5
US-08-700-186-12/c
; Sequence 12, Application US/08700186
; Patent No. 5780286
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick
; APPLICANT: Vockley, Joseph
; TITLE OF INVENTION: ARGINASE II
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,186
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Hap, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50004-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-700-186-12

Query Match 67.0%; Score 13.4; DB 1; Length 43;
Best Local Similarity 93.3%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 actgcagaagcact 18
|||||
DB 16 ACTGCAGAGGCAAT 2

RESULT 6
US-08-914-981-12/c
Sequence 12, Application US/08914981
Patent No. 5912159
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick
APPLICANT: Vockley, Joseph
TITLE OF INVENTION: ARGINASE II
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,981
FILING DATE: 20-AUGUST-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/700,186
FILING DATE: 20-AUGUST-1996
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: ATG50004-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700

TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-914-981-12

Query Match 67.0%; Score 13.4; DB 2; Length 43;
Best Local Similarity 93.3%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 actgcagaagcact 18
|||||
DB 16 ACTGCAGAGGCAAT 2

RESULT 7
US-09-116-115-12/c
Sequence 12, Application US/091161115D
Patent No. 6054308
GENERAL INFORMATION:
APPLICANT: VOCKLEY, JOSEPH G.
APPLICANT: DILLON, PATRICK J.
TITLE OF INVENTION: ARGINASE II
FILE REFERENCE: ATG-50004-3/D1
CURRENT APPLICATION NUMBER: US/09/116,115D
CURRENT FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: US 08/914,981
EARLIER FILING DATE: 1997-08-20
EARLIER APPLICATION NUMBER: US 08/700,186
EARLIER FILING DATE: 1996-08-20
EARLIER APPLICATION NUMBER: US 60/013,395
EARLIER FILING DATE: 1996-03-14
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 43
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-116-115-12

Query Match 67.0%; Score 13.4; DB 3; Length 43;
Best Local Similarity 93.3%; Pred. No. 5e+02;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 actgcagaagcact 18
|||||
DB 16 ACTGCAGAGGCAAT 2

RESULT 8
US-07-825-959-13
Sequence 13, Application US/07825959
Patent No. 5372929
GENERAL INFORMATION:
APPLICANT: Cimino, George C.
APPLICANT: Lin, Lilly
TITLE OF INVENTION: METHOD FOR MEASURING THE INACTIVATION OF
TITLE OF INVENTION: PATHOGENS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Limbach & Limbach
STREET: 2001 Ferry Building

CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/825,959
FILING DATE: 19920127
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Weseman, James C
REGISTRATION NUMBER: 30,507
REFERENCE/DOCKET NUMBER: HRI-02200
TELEPHONE: (415) 433-4150
TELEFAX: (415) 433-8716
TELEX: 278356
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FRAGMENT TYPE: internal
US-07-825-959-13

Query Match 66.0%; Score 13.2; DB 1; Length 27;
Best Local Similarity 83.3%; Pred. No. 5,7e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 aactgcagaagcactga 20
|||||
Db 2 AACTGCAGAAGGCTAGGA 19

RESULT 9
US-08-131-324-13
Sequence 13, Application US/08131324
Patent No. 5565320
GENERAL INFORMATION:
APPLICANT: Cimino, George C.
APPLICANT: Lin, Lily
TITLE OF INVENTION: COMPOUNDS FOR THE PHOTODECONTAMINATION
TITLE OF INVENTION: OF PATHOGENS IN BLOOD
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steritech, Inc.
STREET: 2525 Stanwell Drive
CITY: Concord
STATE: California
COUNTRY: United States of America
ZIP: 94520
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/131,324
FILING DATE: 28-JUN-1993
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US 07/825,959
FILING DATE: 27-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837

REFERENCE/DOCKET NUMBER: STER-1100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 603-9071
TELEFAX: (510) 603-9099
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-131-324-13

Query Match 66.0%; Score 13.2; DB 1; Length 27;
Best Local Similarity 83.3%; Pred. No. 5,7e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 aactgcagaagcactga 20
|||||
Db 2 AACTGCAGAAGGCTAGGA 19

RESULT 10
US-08-377-495-1
Sequence 1, Application US/08377495
Patent No. 5631137
GENERAL INFORMATION:
APPLICANT: Martin, Mark T.
APPLICANT: Smith, Rodger G.
APPLICANT: Darsley, Michael J.
APPLICANT: Simpson, David
APPLICANT: Blackburn, Gary F.
TITLE OF INVENTION: REACTION-BASED SCREENING FOR EXPRESSION
TITLE OF INVENTION: OF AND CONCENTRATION OF CATALYTIC MOIETIES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/377,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/250,934
FILING DATE:
APPLICATION NUMBER: US/08/101,274
FILING DATE:
APPLICATION NUMBER: US/07/841,648
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Barry
REGISTRATION NUMBER: 22,802
REFERENCE/DOCKET NUMBER: 370068-3580
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-377-495-1

Query Match 64.0%; Score 12.8; DB 1; Length 22;

Best Local Similarity 87.5%; Pred. No. 8.6e+02; Indels 0; Gaps 0;

Matches 14; Conservative 0; Mismatches 2;

Qy 1 gaaactgcagaagcca 16

|||||

Db 5 GAAACTGCAGGAGTCA 20

RESULT 11

US-08-748-415-1

; Sequence 1, Application US/08748415

; Patent No. 5891648

; GENERAL INFORMATION:

; APPLICANT: Martin, Mark T.

; APPLICANT: Smith, Rodger G.

; APPLICANT: Darsley, Michael J.

; APPLICANT: Simpson, David

; APPLICANT: Blackburn, Gary F.

; TITLE OF INVENTION: REACTION-BASED SCREENING FOR EXPRESSION

; TITLE OF INVENTION: OF AND CONCENTRATION OF CATALYTIC MOIETIES

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford, P.C.

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: New York

; COUNTRY: United States

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/748.415

; FILING DATE: 13-NOV-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/377,495

; FILING DATE:

; APPLICATION NUMBER: US/08/250,934

; FILING DATE:

; APPLICATION NUMBER: US/08/101,274

; FILING DATE:

; APPLICATION NUMBER: US/07/841,648

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Evans, Barry

; REGISTRATION NUMBER: 22,802

; REFERENCE/DOCKET NUMBER: 370068-3580

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-748-415-1

Query Match

Best Local Similarity 64.0%; Score 12.8; DB 2; Length 22;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gaaactgcagaagcca 16

|||||

Db 5 GAAACTGCAGGAGTCA 20

RESULT 12

US-08-747-654-1

; Sequence 1, Application US/08747654

; Patent No. 6121007

; GENERAL INFORMATION:

; APPLICANT: Martin, Mark T.

; APPLICANT: Smith, Rodger G.

; APPLICANT: Darsley, Michael J.

; APPLICANT: Simpson, David

; APPLICANT: Blackburn, Gary F.

; TITLE OF INVENTION: REACTION-BASED SCREENING FOR EXPRESSION

; TITLE OF INVENTION: OF AND CONCENTRATION OF CATALYTIC MOIETIES

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Curtis, Morris & Safford, P.C.

; STREET: 530 Fifth Avenue

; CITY: New York

; STATE: New York

; COUNTRY: United States

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/747,654

; FILING DATE: 13-NOV-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/377,495

; FILING DATE:

; APPLICATION NUMBER: US/08/250,934

; FILING DATE:

; APPLICATION NUMBER: US/08/101,274

; FILING DATE:

; APPLICATION NUMBER: US/07/841,648

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Evans, Barry

; REGISTRATION NUMBER: 22,802

; REFERENCE/DOCKET NUMBER: 370068-3580

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333

; TELEFAX: (212) 840-0712

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 22 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-747-654-1

Query Match

Best Local Similarity 64.0%; Score 12.8; DB 3; Length 22;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gaaactgcagaagcca 16

|||||

Db 5 GAAACTGCAGGAGTCA 20

RESULT 13

US-09-235-353-1

; Sequence 1, Application US/09235353

; Patent No. 6177270

; GENERAL INFORMATION:

; APPLICANT: Martin, Mark T.

APPLICANT: Smith, Rodger G.
APPLICANT: Darsley, Michael J.
APPLICANT: Simpson, David
APPLICANT: Blackburn, Gary F.
TITLE OF INVENTION: REACTION-BASED SCREENING FOR EXPRESSION
OF AND CONCENTRATION OF CATALYTIC MOIETIES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/235,353
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/748,415
FILING DATE: 13-NOV-1996
APPLICATION NUMBER: US/08/377,495
FILING DATE:
APPLICATION NUMBER: US/08/250,934
FILING DATE:
APPLICATION NUMBER: US/08/101,274
FILING DATE:
APPLICATION NUMBER: US/07/841,648
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Evans, Barry
REGISTRATION NUMBER: 22,802
REFERENCE/DOCKET NUMBER: 370068-3580
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-235-353-1

Query Match 64.0%; Score 12.8; DB 4; Length 22;
Best Local Similarity 87.5%; Pred. No. 8.6e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gaaactgcagaagcca 16
||| ||||| |||
Db 5 GAACTGCAGGAGTCA 20

RESULT 14
US-09-232-477-1
Sequence 1, Application US/09232477
Patent No. 6228846
GENERAL INFORMATION:
APPLICANT: AUDONNET, Jean-Christophe
APPLICANT: BOUCHARDON, Annabelle
APPLICANT: RIVIERE, Michel
TITLE OF INVENTION: POLYNUCLEOTIDE VACCINE FORMULA AGAINST CANINE
PATHOLOGIES, IN PARTICULAR RESPIRATORY AND DIGESTIVE
TITLE OF INVENTION: PATHOLOGIES
FILE REFERENCE: 454313-2240
CURRENT APPLICATION NUMBER: US/09/232,477

CURRENT FILING DATE: 1999-01-15
EARLIER APPLICATION NUMBER: PCT/FR97/01316
EARLIER FILING DATE: 1997-07-15
EARLIER APPLICATION NUMBER: 96/09401
EARLIER FILING DATE: 1996-07-19
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 35
TYPE: DNA
ORGANISM: Borrelia burgdorferi
US-09-232-477-1

Query Match 64.0%; Score 12.8; DB 4; Length 35;
Best Local Similarity 87.5%; Pred. No. 9.2e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 aaactgcagaagccac 17
||||| ||||| |||
Db 2 aaactgcagaatgctc 17

RESULT 15
US-08-105-483-148
Sequence 148, Application US/08105483
Patent No. 5494807
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE
TITLE OF INVENTION: STRAIN
NUMBER OF SEQUENCES: 462
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
ADDRESSEE: c/o William S. Frommer
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/105,483
FILING DATE: 12-AUG-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,951
FILING DATE: 06-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2400
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-105-483-148

Query Match 63.0%; Score 12.6; DB 1; Length 21;
Best Local Similarity 78.9%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 gaaactgcagaagccactg 19

Db 3 GAAAGAGCAGAGACAGTG 21

Search completed: October 2, 2001, 16:03:46
Job time: 14590 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:44 ; Search time 417.38 Seconds
(without alignments)
9.071 Million cell updates/sec

Title: US-09-757-100B-4
Perfect score: 20
Sequence: 1 ggcgcgtgaagcgaaggca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	3	US-09-377-310-4
2	15	75.0	15	3	US-09-377-310-24
3	13.2	66.0	21	1	US-08-250-849-19
4	13.2	66.0	21	1	US-08-434-474-19
5	12.8	64.0	20	4	US-09-180-437-51
6	12.8	64.0	20	4	US-09-180-437-52
7	12.8	64.0	20	4	US-09-180-437-53
8	12.2	61.0	17	3	US-08-598-099-6
9	12.2	61.0	38	1	US-08-750-077-4
10	12.2	61.0	40	3	US-08-841-267-15
11	12.2	61.0	50	2	US-08-472-171-46
12	12.2	61.0	50	2	US-08-894-526-46
13	12.2	61.0	50	3	US-09-013-047-46
14	12.2	61.0	50	3	US-09-374-597-46
15	12	60.0	20	1	US-07-678-448A-4
16	12	60.0	20	4	US-09-560-594-41
17	12	60.0	27	3	US-09-006-597-16
18	11.8	59.0	15	4	US-09-180-437-217
19	11.8	59.0	20	4	US-09-180-437-50
20	11.6	58.0	27	1	US-08-624-545-32
21	11.6	58.0	43	1	US-08-428-733A-46
22	11.4	57.0	20	4	US-09-517-584A-22
23	11.4	57.0	20	4	US-09-180-437-48
24	11.4	57.0	20	4	US-09-180-437-49
25	11.2	56.0	17	3	US-08-998-099-5
26	11.2	56.0	20	3	US-09-280-799-185
27	11.2	56.0	25	5	PCT-US92-02977-4

28	11.2	56.0	25	5	PCT-US95-03032-7	Sequence 7, Appli
c 29	11.2	56.0	30	1	US-08-015-770B-72	Sequence 72, Appl
c 30	11.2	56.0	38	1	US-08-399-580B-11	Sequence 11, Appl
31	11.2	56.0	33	1	US-08-438-639-16	Sequence 16, Appl
32	11.2	56.0	33	1	US-07-813-338A-16	Sequence 16, Appl
33	11.2	56.0	33	3	US-08-441-971-91	Sequence 91, Appl
34	11.2	56.0	33	4	US-08-221-653-91	Sequence 91, Appl
35	11.2	56.0	33	4	US-08-442-144A-91	Sequence 91, Appl
36	11.2	56.0	37	2	US-08-570-155-8	Sequence 8, Appli
37	11.2	56.0	37	5	PCT-US95-02861-8	Sequence 8, Appli
c 38	11.2	56.0	38	1	US-07-841-662-21	Sequence 21, Appl
c 39	11.2	56.0	38	1	US-08-209-797-21	Sequence 21, Appl
c 40	11.2	56.0	38	1	US-08-669-685-21	Sequence 21, Appl
c 41	11.2	56.0	38	3	US-09-103-486-21	Sequence 21, Appl
c 42	11.2	56.0	38	5	PCT-US93-01557-21	Sequence 21, Appl
c 43	11.2	56.0	41	1	US-07-951-715A-87	Sequence 87, Appl
44	11.2	56.0	41	2	US-08-459-448A-87	Sequence 87, Appl
45	11.2	56.0	41	3	US-08-459-595A-87	Sequence 87, Appl

ALIGNMENTS

RESULT 1
US-09-377-310-4
; Sequence 4, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-4

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ggcgcgtgaagcgaaggca 20
|||||
Db 1 ggcgcgtgaagcgaaggca 20

RESULT 2
US-09-377-310-24
; Sequence 24, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-24

Query Match 75.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 cgcctgaagcgaag 17
|||||
Db 1 cgcctgaagcgaag 15

RESULT 3
US-08-250-849-19
; Sequence 19, Application US/08250849
; Patent No. 5567583
; GENERAL INFORMATION:
; APPLICANT: Chang-Ning J. Wang and Kai-
; APPLICANT: Yuan Wu
; TITLE OF INVENTION: METHOD FOR DETECTING A TARGET
; TITLE OF INVENTION: NUCLEIC ACID
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/250,849
; FILING DATE: 05/26/94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/808,463
; FILING DATE: December 16, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Y. Rocky Tsao
; REGISTRATION NUMBER: 34,053
; REFERENCE/DOCKET NUMBER: 06498/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-250-849-19

Query Match 66.0%; Score 13.2; DB 1; Length 21;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 cgcctgaagcgaagc 19
|||||
Db 1 GCGGCGTGAAGCGCGC 18

RESULT 4
US-08-434-474-19
; Sequence 19, Application US/08434474
; Patent No. 5712386

; GENERAL INFORMATION:
; APPLICANT: Wang et al.
; TITLE OF INVENTION: METHOD FOR DETECTING A TARGET
; TITLE OF INVENTION: NUCLEIC ACID
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,474
; FILING DATE: 05/04/95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/250,849
; FILING DATE: 05/26/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Y. Rocky Tsao
; REGISTRATION NUMBER: 34,053
; REFERENCE/DOCKET NUMBER: 06498/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-434-474-19

Query Match 66.0%; Score 13.2; DB 1; Length 21;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 cgcctgaagcgaagc 19
|||||
Db 1 GCGGCGTGAAGCGCGC 18

RESULT 5
US-09-180-437-51
; Sequence 51, Application US/09180437
; Patent No. 6251873
; GENERAL INFORMATION:
; APPLICANT: FUKUSAKO, Shioji
; APPLICANT: MORISAWA, Yoshifumi
; APPLICANT: KUSUYAMA, Takeshi
; TITLE OF INVENTION: Antisense Compounds to CD14
; FILE REFERENCE: 1110-209P
; CURRENT APPLICATION NUMBER: US/09/180,437
; CURRENT FILING DATE: 1998-11-06
; EARLIER APPLICATION NUMBER: PCT/JP98/00953
; EARLIER FILING DATE: 1998-03-09
; EARLIER APPLICATION NUMBER: 09-053518 JAPAN
; EARLIER FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 51
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: other nucleic

; OTHER INFORMATION: acid, synthetic DNA
US-09-180-437-51

Query Match 64.0%; Score 12.8; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 6.1e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 ccgtgaagcgaaggca 20
|| ||||| |||||
Db 1 ccctgaagccaaggca 16

RESULT 6

US-09-180-437-52

; Sequence 52, Application US/09180437
; Patent No. 6251873
; GENERAL INFORMATION:
; APPLICANT: FUKUSAKO, Shioji
; APPLICANT: MORISAWA, Yoshifumi
; APPLICANT: KUSUYAMA, Takeshi
; TITLE OF INVENTION: Antisense Compounds to CD14
; FILE REFERENCE: 1110-209P
; CURRENT APPLICATION NUMBER: US/09/180,437
; CURRENT FILING DATE: 1998-11-06
; EARLIER APPLICATION NUMBER: PCT/JP98/00953
; EARLIER FILING DATE: 1998-03-09
; EARLIER APPLICATION NUMBER: 09-053518 JAPAN
; EARLIER FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:other nucleic
; OTHER INFORMATION: acid, synthetic DNA
US-09-180-437-52

Query Match 64.0%; Score 12.8; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 6.1e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 ccgtgaagcgaaggca 20
|| ||||| |||||
Db 1 ccctgaagccaaggca 16

RESULT 7

US-09-180-437-53

; Sequence 53, Application US/09180437
; Patent No. 6251873
; GENERAL INFORMATION:
; APPLICANT: FUKUSAKO, Shioji
; APPLICANT: MORISAWA, Yoshifumi
; APPLICANT: KUSUYAMA, Takeshi
; TITLE OF INVENTION: Antisense Compounds to CD14
; FILE REFERENCE: 1110-209P
; CURRENT APPLICATION NUMBER: US/09/180,437
; CURRENT FILING DATE: 1998-11-06
; EARLIER APPLICATION NUMBER: PCT/JP98/00953
; EARLIER FILING DATE: 1998-03-09
; EARLIER APPLICATION NUMBER: 09-053518 JAPAN
; EARLIER FILING DATE: 1997-03-07
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:other nucleic
; OTHER INFORMATION: acid, synthetic DNA
US-09-180-437-53

Query Match 64.0%; Score 12.8; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 6.1e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 5 ccgtgaagcgaaggca 20
|| ||||| |||||
Db 4 ccctgaagccaaggca 19

RESULT 8

US-08-998-099-6/C

; Sequence 6, Application US/08998099A
; Patent No. 6103890
; GENERAL INFORMATION:
; APPLICANT: JARVIS, THALE
; APPLICANT: MCSWIGGEN, JAMES A.
; APPLICANT: STINCHCOMB, DAN T.
; TITLE OF INVENTION: ENZYMTIC NUCLEIC ACID TREATMENT OF DISEASES
; TITLE OF INVENTION: OR CONDITIONS RELATED TO LEVELS OF C-FOS
; FILE REFERENCE: 231/175
; CURRENT APPLICATION NUMBER: US/08/998,099A
; CURRENT FILING DATE: 1997-12-24
; EARLIER APPLICATION NUMBER: 60/037,658
; EARLIER FILING DATE: 1997-01-23
; EARLIER APPLICATION NUMBER: 08/373,124
; EARLIER FILING DATE: 1995-01-13
; EARLIER APPLICATION NUMBER: 08/245,466
; EARLIER FILING DATE: 1994-05-18
; NUMBER OF SEQ ID NOS: 375
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-08-998-099-6

Query Match 61.0%; Score 12.2; DB 3; Length 17;
Best Local Similarity 82.4%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ggcgcgcgtgaagcgaag 17
|| ||||| |||||
Db 17 GGCCTGTGAAGCAGAG 1

RESULT 9

US-08-750-077-4

; Sequence 4, Application US/08750077
; Patent No. 5783423
; GENERAL INFORMATION:
; APPLICANT: Wood, Patricia C
; APPLICANT: Quirk, Alan V
; TITLE OF INVENTION: Yeast Strains
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon LLC
; STREET: 1020 First Avenue
; CITY: King of Prussia
; COUNTRY: USA
; ZIP: PA 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,077

```
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411356.0
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/01317
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Moore, Steven J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 878 4073
; TELEFAX: (610) 878 4221
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..38
; OTHER INFORMATION: /note="Oligonucleotide for PCR
; OTHER INFORMATION: amplification of the 3' end of the Hsp150 gene."
US-08-750-077-4

Query Match 61.0%; Score 12.2; DB 1; Length 38;
Best Local Similarity 82.4%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 gccgtgaagcgaagca 20
Db 1 GCCGTGACGAGGGA 17

RESULT 10
US-08-841-267-15
; Sequence 15, Application US/08841267C
; Patent No. 6007990
; GENERAL INFORMATION:
; APPLICANT: Levine, Robert A.
; TITLE OF INVENTION: Detection and Quantification of One or
; TITLE OF INVENTION: More Nucleotide Sequence Target Analytes in a Sample Using
; TITLE OF INVENTION: More Nucleotide Sequence Target Analytes in a Sample Using
; TITLE OF INVENTION: Spatially Localized Target Analyte Replication
; FILE REFERENCE: UFB-001
; CURRENT APPLICATION NUMBER: US/08/841.267C
; CURRENT FILING DATE: 1997-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 40
; TYPE: DNA
; ORGANISM: virus HBV
US-08-841-267-15

Query Match 61.0%; Score 12.2; DB 3; Length 40;
Best Local Similarity 82.4%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggcgcgcgtgaagcgaag 17
Db 16 gacacggtgaagcgaag 32

RESULT 11
US-08-472-171-46
; Sequence 46, Application US/08472171
```

```
; Patent No. 5932714
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yacoub, Reza K.
; APPLICANT: Zealey, Gavin R.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Expression Of Gene Products From
; TITLE OF INVENTION: Genetically Manipulated Strains Of Bordetella
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, Suite 701
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,171
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/393,334
; FILING DATE: 23-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-507 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 416-595-1155
; TELEFAX: 416-595-1163
; TELEX: 065-24567 Simbas
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-171-46

Query Match 61.0%; Score 12.2; DB 2; Length 50;
Best Local Similarity 82.4%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ggcgcgcgtgaagcgaag 17
Db 23 GCGCGCGTGACGCGCAAG 39

RESULT 12
US-08-894-526-46
; Sequence 46, Application US/08894526
; Patent No. 5942418
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Yacoub, Reza K
; APPLICANT: Zealey, Gavin R
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: EXPRESSION OF GENE PRODUCTS FROM
; TITLE OF INVENTION: GENETICALLY MANIPULATED STRAINS OF BORDETELLA
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,526
FILING DATE: 01-DEC-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-724 MTS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-894-526-46

Query Match 61.0%; Score 12.2; DB 2; Length 50;
Best Local Similarity 82.4%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggccgcgtgaagcgaag 17
||||| | | | |
Db 23 GGC GCGT GACG CCAAG 39

RESULT 13

US-09-013-047-46
Sequence 46, Application US/09013047
Patent No. 598168
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
APPLICANT: Yacoub, Reza K.
APPLICANT: Zealey, Gavin H.
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Expression Of Gene Products From
Genetically Manipulated Strains Of Bordetella
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,047
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,171
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/393,334
FILING DATE: 23-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-507 MTS:vg
TELECOMMUNICATION INFORMATION:

TELEPHONE: 416-595-1155
TELEFAX: 416-595-1163
TELEX: 065-24567 Simbas
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-013-047-46

Query Match 61.0%; Score 12.2; DB 2; Length 50;
Best Local Similarity 82.4%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggccgcgtgaagcgaag 17
||||| | | | |
Db 23 GGC GCGT GACG CCAAG 39

RESULT 14

US-09-374-597-46
Sequence 46, Application US/09374597
Patent No. 6140082
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena M.
APPLICANT: Yacoub, Reza K.
APPLICANT: Zealey, Gavin H.
APPLICANT: Klein, Michel H.
TITLE OF INVENTION: Expression Of Gene Products From
Genetically Manipulated Strains Of Bordetella
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/374,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/393,334
FILING DATE: FEBRUARY 23, 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I.
REGISTRATION NUMBER: 24973
REFERENCE/DOCKET NUMBER: 1038-964
TELECOMMUNICATION INFORMATION:
TELEPHONE: 416-595-1155
TELEFAX: 416-595-1163
TELEX: 065-24567 Simbas
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-374-597-46

Query Match 61.0%; Score 12.2; DB 3; Length 50;
Best Local Similarity 82.4%; Pred. No. 1.2e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggcgcgtgaagcgaag 17
|||||
Db 23 GCGCGGTGACGCAAG 39
RESULT 15
US-07-678-448A-4/c
; Sequence 4, Application US/07678448A
; Patent No. 5521301
; GENERAL INFORMATION:
; APPLICANT: R. Bruce Wallace
; APPLICANT: Luis Ugozzoli
; TITLE OF INVENTION: Genotyping of Multiple Allele
; TITLE OF INVENTION: Systems
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: City of Hope
; STREET: 1500 East Duarte Road
; CITY: Duarte
; STATE: California
; COUNTRY: United States of America
; ZIP: 91010-0269
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3M Double Density 5 1/4" diskette
; COMPUTER: Wang PC
; OPERATING SYSTEM: MS-DOS (R) Version 3.30
; SOFTWARE: Microsoft (R)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/678,448A
; FILING DATE: 19910401
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/283,142
; FILING DATE: 12 December, 1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Irons, Edward S.
; REGISTRATION NUMBER: 16,541
; REFERENCE/DOCKET NUMBER: NO. 5521301e
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 785-6938
; TELEFAX: (202) 785-5351
; TELEX: 440087 LM WSH
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20
; TYPE: NUCLEIC ACID
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: No. 5521301 Applicable
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: No. 5521301 Applicable
; ORIGINAL SOURCE: Synthetically Prepared
; IMMEDIATE SOURCE: Synthetically Prepared
; POSITION IN GENOME: No. 5521301e
; FEATURE: No. 5521301e
; PUBLICATION INFORMATION: No. 5521301e
US-07-678-448A-4

Query Match 60.0%; Score 12; DB 1; Length 20;
Best Local Similarity 75.0%; Pred. NO. 1.5e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ggcgcgtgaagcgaagca 20
|||||
Db 20 GCGCCTACAGCGCTGGCA 1

Search completed: October 2, 2001, 16:03:45
Job time: 14589 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:33 ; Search time 876.95 seconds
(without alignments)
14.320 Million cell updates/sec

Title: US-09-757-100B-4

Perfect score: 20

Sequence: 1 ggcgcgtgaagcgaaggca 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	22	AAC65536 Human focal adhesi
2	15	75.0	15	22	AAC65556 Human focal adhesi
3	14.2	71.0	40	18	AAT87270 IL-4 2'F/NH2 RNA 1
4	14.2	71.0	40	18	AAT87201 Interleukin-4 2'F
5	13.2	66.0	21	17	AAT09720 Human AML-1 limit
6	13.2	66.0	21	18	AAT45484 Human leukaemia br
7	13.2	66.0	21	19	AAV09361 Limiting primer us
8	12.8	64.0	21	22	AAE95308 Human gene single
9	12.6	63.0	33	21	AA443694 Mass spectrometric
10	12.6	63.0	36	19	AAV36074 Oligonucleotide SC
11	12.6	63.0	50	21	AAA96608 Nucleotide sequenc

c	12	12.2	61.0	17	19	AAV95266 Human c-fos target
c	13	12.2	61.0	19	21	AAA84390 Cyclin D3 ribozyme
c	14	12.2	61.0	20	22	AAF74126 Primer #60. Homo
c	15	12.2	61.0	22	22	AAF69762 Human IL4Ra1pha ge
c	16	12.2	61.0	29	21	AAA04651 Polymorphic fragme
c	17	12.2	61.0	31	13	AAQ25255 NANBH PCR primer p
c	18	12.2	61.0	38	17	AAT07272 Primer for HBV pX
c	19	12.2	61.0	40	19	AAV08311 Primer for HBV pX
c	20	12.2	61.0	50	10	AAV91945 Complementaty stra
c	21	12	60.0	20	17	AAT30041 Allele-specific PC
c	22	12	60.0	21	22	AAF23962 Bacillus lichenifo
c	23	12	60.0	27	19	AAV07603 Nucleotide sequenc
c	24	12	60.0	36	21	AAA35716 permutein linker e
c	25	12	60.0	38	13	AAQ25438 Probe for alkaline
c	26	12	60.0	38	17	AAT34327 Alkaline protease
c	27	12	60.0	50	21	AAV77403 Human clone cg4492
c	28	11.8	59.0	20	20	AAZ00590 Human GPC4 exon 2
c	29	11.8	59.0	30	22	AAF73589 HGF nucleic acid 1
c	30	11.8	59.0	42	20	AAQ21515 Integrin alpha 6 s
c	31	11.6	58.0	22	21	AAQ01060 Oligo #5 for site-
c	32	11.6	58.0	24	21	AAC58525 Human PRO1434 (UNQ
c	33	11.6	58.0	24	21	AAV51273 Forward primer for
c	34	11.6	58.0	25	21	AAA30072 PCR primer 34387.t
c	35	11.6	58.0	25	22	AAF60392 PRO240 probe #2.
c	36	11.6	58.0	27	21	AAAG0931 Phanochoaete sord
c	37	11.6	58.0	34	21	AAZ43239 PCR primer for C.
c	38	11.6	58.0	35	14	AAQ51295 D. nodosus proteas
c	39	11.6	58.0	35	20	AAV31843 Porphyromonas ging
c	40	11.6	58.0	36	18	AAAT69182 Oligonucleotide Kb
c	41	11.6	58.0	38	20	AAV78995 HIF-1alpha gene am
c	42	11.6	58.0	43	16	AAT02527 TnI gene forward p
c	43	11.6	58.0	43	19	AAV32360 Human troponin I f
c	44	11.6	58.0	43	20	AAV25907 Human troponin I g
c	45	11.6	58.0	48	18	AAT97348 Mouse CGRP recepto

ALIGNMENTS

RESULT 1
AAC65536
ID AAC65536 standard; DNA; 20 BP.
XX
AC AAC65536;
XX
DT 12-FEB-2001 (first entry)
XX
DE Human focal adhesion kinase antisense sequence #2.
XX
KW Human; focal adhesion kinase; FAK; signal transduction; cancer;
KW embryonic development disorder; angiogenic disorder; wound healing;
KW antisense; phosphothioate; ss.
XX
OS Homo sapiens.
XX
PN US6133031-A.
XX
PD 17-OCT-2000.
XX
PF 19-AUG-1999; 99US-0377310.
XX
PR 19-AUG-1999; 99US-0377310.
XX
PA (ISIS-) ISIS PHARM INC.
PI Monia BP, Gaarde WA;
XX
DR WPI; 2001-006141/01.
XX
PT New antisense compounds for inhibiting focal adhesion kinase
PT expression, especially useful for inhibiting retinal
PT neovascularization, or for diagnosing and treating e.g. colon cancer -

PS Claim 3; Column 23; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
CC sequences to the human focal adhesion kinase (FAK) protein. This protein
CC is involved in integrin-mediated signal transduction, and is implicated
CC in cancer, particularly colon, breast and oral tumours, embryonic
CC development disorders, angiogenic disorders and wound healing. The
CC antisense sequences, including the one shown here, can be used in the
CC treatment of all of these.

XX Sequence 20 BP; 5 A; 5 C; 9 G; 1 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.66;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ggcgcgtgaagcgaagca 20

|||||

Db 1 ggcgcgtgaagcgaagca 20

RESULT 2

AAC65556

ID AAC65556 standard; DNA; 15 BP.

AC AAC65556;

DT 12-FEB-2001 (first entry)

XX

DE Human focal adhesion kinase antisense sequence #22.

XX

Human; focal adhesion kinase; FAK; signal transduction; cancer;

KW embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphorothioate; ss.

XX

OS Homo sapiens.

XX

PN US6133031-A.

XX

PD 17-OCT-2000.

XX

PF 19-AUG-1999; 99US-0377310.

XX

PR 19-AUG-1999; 99US-0377310.

XX

PA (ISIS-) ISIS PHARM INC.

XX

PI Monia BP, Gaarde WA;

XX

WPI; 2001-006141/01.

XX

PT New antisense compounds for inhibiting focal adhesion kinase

PT expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

XX

PS Example 2; Column 25; 30pp; English.

XX

CC The present invention describes a number of phosphorothioate antisense
CC sequences to the human focal adhesion kinase (FAK) protein. This protein
CC is involved in integrin-mediated signal transduction, and is implicated
CC in cancer, particularly colon, breast and oral tumours, embryonic
CC development disorders, angiogenic disorders and wound healing. The
CC antisense sequences, including the one shown here, can be used in the
CC treatment of all of these.

XX

SQ Sequence 15 BP; 4 A; 4 C; 6 G; 1 T; 0 other;

Query Match

Best Local Similarity 75.0%; Score 15; DB 22; Length 15;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 cgccgtgaagcgaag 17

|||||

Db 1 cgccgtgaagcgaag 15

RESULT 3

AAT87270

ID AAT87270 standard; RNA; 40 BP.

XX

AC AAT87270;

XX

DT 16-NOV-1997 (first entry)

XX

DE IL-4 2'/NH2 RNA ligand, nitrocellulose filter binding clone #21.

XX

high affinity ligand; cytokine; interferon-gamma; IFN-gamma; RANTES;

KW interleukin-4; IL-4; tumour necrosis factor-alpha; TNF-alpha;

KW Systematic Evolution of Ligands by Exponential enrichment; SELEX;

KW diagnosis; inflammatory response; septic shock; arthritis;

KW graft-vs-host reaction; ss.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT modified_base 1..40

FT /*tag= a

FT /note= "all U's are 2'-NH2 modified and all C's are

FT 2'-F modified"

XX

PN WO9640717-A1.

XX

PD 19-DEC-1996.

XX

PF 04-JUN-1996; 96WO-US09537.

XX

PR 07-JUN-1995; 95US-0481710.

XX

PR 07-JUN-1995; 95US-0477527.

XX

PA (NEXS-) NEXSTAR PHARM INC.

XX

PI Gold L, Jayasena S, Pagratis N, Tasset D;

XX

WPI; 1997-087049/08.

XX

PT Identification of nucleic acid ligands that bind cytokine(s) - by

PT partitioning the ligands from a nucleic acid mixture, using SELEX

PT techniques

XX

PS Claim 20; Page 69; 175pp; English.

XX

AAT87178-276 are interleukin (IL)-4 RNA ligands obtained after 17 rounds

CC of SELEX (Systematic Evolution of Ligands by Exponential enrichment).

CC The sequences were analysed for conserved sequences and aligned by this

CC criterion. The 2'F sequences fell into a single group with 12 orphan

CC sequences. 2'NH2 sequences fell into 2 distinct groups. Group 1 were

CC shown to bind to IL-4 and the other group were shown to bind to

CC nitrocellulose filters and were identified by the presence of a direct

CC repeat of the sequence GGAGG. A single orphan 2'NH2 sequence was also

CC found. The 2'F/NH2 sequences fell into 3 groups, of which one group

CC bound to nitrocellulose filters. Nucleic acid (NA) ligands to a cytokine

CC can be identified using SELEX. A candidate mixture of NA's are contacted

CC with a cytokine where the NA's having an increased affinity to the

CC cytokine relative to the candidate mixture may be partitioned from the

CC rest of the mixture. The NA's with increased affinity are amplified to

CC yield a mixture of NA's enriched for the NA sequences and relatively

CC higher affinity and selectivity for binding to the cytokine. The NA

CC ligands are useful in diagnostic and therapeutic applications especially

CC to prevent or treat diseases or medical conditions in human patients,

CC e.g. associated with excessive cytokine production such as inflammatory

CC responses mediated by IFN-gamma or interleukin-4, septic shock,

CC arthritis or graft-vs-host reactions mediated by tumour necrosis

CC factor-alpha.

XX

SQ Sequence 40 BP; 12 A; 5 C; 18 G; 5 U; 0 other;

Query Match 71.0%; Score 14.2; DB 18; Length 40;
Best Local Similarity 78.9%; Pred. No. 4.1e+02;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 gcgcgtgaagcgaagca 20
|||||:|||||
DB 5 gcgccaugaagcaagga 23

RESULT 4

AAT87201
ID AAT87201 standard; RNA; 40 BP.

XX AC AAT87201;

XX DT 08-NOV-1997 (first entry)

XX DE Interleukin-4 2'F RNA ligand, orphan clone #21.

XX KW high affinity ligand; cytokine; interferon-gamma; IFN-gamma; RANTES;
KW Interleukin-4; IL-4; tumour necrosis factor-alpha; TNF-alpha;
KW Systematic Evolution of Ligands by Exponential enrichment; SELEX;
KW diagnosis; inflammatory response; septic shock; arthritis;
KW graft-vs-host reaction; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT modified_base 1..40

FT /*tag= a
FT /note= "all pyrimidines are 2'-F"

XX PN W09640717-A1.

XX PD 19-DEC-1996.

XX PF 04-JUN-1996; 96WO-US09537.

XX PR 07-JUN-1995; 95US-0481710.

XX PR 07-JUN-1995; 95US-0477527.

XX PA (NEXS-) NEXSTAR PHARM INC.

XX PI Gold L, Jayasena S, Pagratis N, Tasset D;

XX DR WPI; 1997-087049/08.

XX PT Identification of nucleic acid ligands that bind cytokine(s) - by
PT partitioning the ligands from a nucleic acid mixture, using SELEX
PT techniques

XX PS Claim 20; Page 65; 175pp; English.

XX AAT87178-276 are interleukin (IL)-4 RNA ligands obtained after 17 rounds
CC of SELEX (Systematic Evolution of Ligands by Exponential enrichment).
CC The sequences were analysed for conserved sequences and aligned by this
CC criterion. The 2'F sequences fell into a single group with 12 orphan
CC sequences. 2'NH2 sequences fell into 2 distinct groups. Group 1 were
CC shown to bind to IL-4 and the other group were shown to bind to
CC nitrocellulose filters and were identified by the presence of a direct
CC repeat of the sequence GGAGG. A single orphan 2'NH2 sequence was also
CC found. The 2'F/NH2 sequences fell into 3 groups, of which one group
CC bound to nitrocellulose filters. Nucleic acid (NA) ligands to a cytokine
CC can be identified using SELEX. A candidate mixture of NA's are contacted
CC with a cytokine where the NA's having an increased affinity to the
CC cytokine relative to the candidate mixture may be partitioned from the
CC rest of the mixture. The NA's with increased affinity are amplified to
CC yield a mixture of NA's enriched for the NA sequences and relatively
CC higher affinity and selectivity for binding to the cytokine. The NA
CC ligands are useful in diagnostic and therapeutic applications especially

CC to prevent or treat diseases or medical conditions in human patients,
CC e.g. associated with excessive cytokine production such as inflammatory
CC responses mediated by IFN-gamma or interleukin-4, septic shock,
CC arthritis or graft-vs-host reactions mediated by tumour necrosis
CC factor-alpha.

SQ Sequence 40 BP; 12 A; 5 C; 18 G; 5 U; 0 other;

Query Match 71.0%; Score 14.2; DB 18; Length 40;
Best Local Similarity 78.9%; Pred. No. 4.1e+02;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 gcgcgtgaagcgaagca 20
|||||:|||||
DB 5 gcgccaugaagcaagga 23

RESULT 5

AAT09720
ID AAT09720 standard; DNA; 21 BP.

XX AC AAT09720;

XX DT 27-JUN-1996 (first entry)

XX DE Human AML-1 limiting primer, for asymmetric amplification.

XX KW Polymerase chain reaction; amplification; non-specific priming;
KW blocking oligonucleotide; donor; acceptor; fluorophore;
KW energy transfer; ligation; AML-1; acute myeloid leukaemia;
KW breakpoint related sequence; ss.

XX OS Synthetic.

XX PN W09532306-A1.

XX PD 30-NOV-1995.

XX PF 23-MAY-1994; 94WO-US05767.

XX PR 23-MAY-1994; 94WO-US05767.

XX PA (BIOT-) BIOTRONICS CORP.

XX PI Wang CJ, Wu K;

XX DR WPI; 1996-020598/02.

XX PT Detecting target nucleic acid by amplification - with primer-
PT blocking oligo:nucleotide duplex(es) labelled with donor and
PT acceptor fluorophore(s), to reduce non-specific priming

XX PS Example 4; Page 23; 41pp; English.

XX The presence of a blocking oligonucleotide partially complementary
CC to an amplification primer in a PCR mixture reduces the number of
CC non-specific priming events. When labelled with a fluorophore, the
CC blocking oligonucleotide can also be used to monitor the amplification
CC process by participating in fluorescence energy transfer. This
CC energy transfer can be enhanced by using the blocking oligonucleotide
CC as a template for ligation of its complementary sequence to the
CC primer. In an example, the human acute myeloid leukaemia breakpoint
CC related sequence was asymmetrically amplified using an excess primer
CC and a limiting primer (see AAT09719 and AAT09720, respectively). The
CC amplification process could be monitored using either a primer:
CC blocking oligonucleotide duplex or a universal detection duplex
CC coupled to a primer.

SQ Sequence 21 BP; 2 A; 6 C; 11 G; 2 T; 0 other;

Query Match

66.0%; Score 13.2; DB 17; Length 21;

Best Local Similarity 83.3%; Pred. No. 1.2e+03; Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gcgcgtgaagcgagc 19
 ||| ||||| |||
 Db 1 gcgcgtgaagcgagc 18

RESULT 6
 AAT45484
 ID AAT45484 standard; DNA; 21 BP.
 AC AAT45484;
 XX
 DT 02-APR-1997 (first entry)
 XX
 DE Human leukaemia breakpoint related sequence, limiting primer.
 XX
 KW Detection duplex; asymmetric amplification; human; acute; myeloid;
 KW leukaemia; breakpoint; related sequence; AMP-1; X chromosome;
 KW specific amelogenin; AMG-X; primer; excess; limiting; ss.
 OS Synthetic.
 XX
 PN US5567383-A.
 PD 22-OCT-1996.
 XX
 PF 16-DEC-1991; 91US-0808463.
 XX
 PR 26-MAY-1994; 94US-0250849.
 PR 16-DEC-1991; 91US-0808463.
 XX
 PA (BIOT-) BIOTRONICS CORP.
 XX
 PI Wang CJ, Wu K;
 XX WPI; 1997-010704/01.
 DR
 XX Nucleic acid detection by amplification - in presence of
 PT primer-complementary oligo:nucleotide to block non-specific priming
 XX
 PS Example 4; Column 12; 17pp; English.
 CC In an example to demonstrate the application of a detection duplex
 CC to monitor the asymmetric amplification of target nucleic acids,
 CC the human acute myeloid leukaemia breakpoint related sequence,
 CC (AMP-1), and the human X chromosome specific amelogenin (AMG-X)
 CC were used as amplification targets. The sequences of the primers
 CC duplex comprised AAT45481 and AAT45482 and the sequences of
 CC excess and limiting primers for AML-1 and AMG-X respectively were
 CC AAT45483 and AAT45484, and AAT45485 and AAT45486. Male human genomic DNA
 CC sequences were asymmetrically amplified with excess primer for
 CC 20 cycles prior to the addition of the primer duplex. The results
 CC compiled from detection at cycle 30 demonstrated the quantitative
 CC resolution and detection sensitivity by the duplex. The decrease
 CC in fluorescence intensity reflected the initial target doses.
 XX
 SQ Sequence 21 BP; 2 A; 6 C; 11 G; 2 T; 0 other;

Query Match 66.0%; Score 13.2; DB 18; Length 21;
 Best Local Similarity 83.3%; Pred. No. 1.2e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 gcgcgtgaagcgagc 19
 ||| ||||| |||
 Db 1 gcgcgtgaagcgagc 18

RESULT 7
 AAV09361
 ID AAV09361 standard; DNA; 21 BP.
 AC AAV09361;
 XX
 DT 15-MAY-1998 (first entry)
 XX
 DE Limiting primer used for detecting target nucleic acid AML-1.
 XX
 KW Target; DNA duplex; donor; acceptor; fluorescent label; fluorophore;
 KW energy transfer; polymerase; primer; acute myeloid leukaemia; AML-1; ss.
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN US5712386-A.
 PD 27-JAN-1998.
 XX
 PF 04-MAY-1995; 95US-0434474.
 XX
 PR 26-MAY-1994; 94US-0250849.
 PR 16-DEC-1991; 91US-0808463.
 PR 04-MAY-1995; 95US-0434474.
 XX
 PA (BIOT-) BIOTRONICS CORP.
 XX
 PI Wang CJ, Wu K;
 XX WPI; 1998-120033/11.
 DR
 XX Kits for detecting target nucleic acids - and DNA duplexes with
 PT donor and acceptor fluorescent labels
 XX
 PS Example 4; Column 12; 17pp; English.
 XX
 CC This is a limiting primer used in a kit for detecting a target nucleic
 CC acid of human acute myeloid leukaemia breakpoint related sequence
 CC (AML-1). The kit is a DNA duplex which comprises a first oligonucleotide
 CC capable of acting as a primer, with or without a segment noncontiguous to
 CC its priming sequence, for use with a polymerase in the amplification of a
 CC target nucleic acid, a second oligonucleotide which is hybridised, via
 CC at least 5 consecutive fully complementary nucleotide pairings, with the
 CC first oligonucleotide, the second oligonucleotide being incapable of
 CC acting as a primer for the polymerase, and a first fluorophore covalently
 CC attached to the first oligonucleotide, with one of the two fluorophores
 CC being a donor fluorophore and the other being an acceptor fluorophore, so
 CC that when the two fluorophores are in close proximity resonance energy
 CC transfer between them is allowed. Each of the first oligonucleotide and
 CC the second oligonucleotide contains 10--50 nucleotides. Another kit
 CC claimed comprises a first and second primer both optionally having a
 CC segment non-contiguous to a first or second priming sequence,
 CC respectively, which are used with a polymerase for the amplification of
 CC the target nucleic acid and an oligonucleotide which is incapable of
 CC acting as a primer for the polymerase and has at least 5 consecutive
 CC nucleotides fully complementary to at least 5 consecutive nucleotides of
 CC the first primer. Each of the first primer, the second primer and the
 CC oligonucleotide contains 10-50 nucleotides. A third kit for detecting a
 CC target nucleic acid comprises a first oligonucleotide being incapable of
 CC acting as a primer for use with a polymerase in the amplification of a
 CC target nucleic acid, and containing 10-50 nucleotides with a first
 CC fluorophore covalently attached to it, and a second oligonucleotide
 CC containing 5-30 nucleotides with a second fluorophore covalently attached
 CC to it, the second oligonucleotide having a free 3' OH and being capable
 CC of hybridizing, via at least 5 consecutive fully complementary nucleotide
 CC pairings, with the first oligonucleotide. The first oligonucleotide has
 CC an overhang beyond the 3' end of the second oligonucleotide by 1-12
 CC nucleotides, and the first and second oligonucleotides, one of which is a
 CC donor fluorophore and the other an acceptor fluorophore are in close
 CC proximity when the first oligonucleotide hybridises to the second
 CC oligonucleotide to allow resonance energy transfer between them. The
 CC kits are used in homogeneous assays in which the target nucleic acid
 CC sequence is amplified and the amplified target is detected without

CC conducting a separation step.

XX Sequence 21 BP; 2 A; 6 C; 11 G; 2 T; 0 other;

SQ

Query Match 66.0%; Score 13.2; DB 19; Length 21;

Best Local Similarity 83.3%; Pred. No. 1.2e+03;

Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 gcgcgtgaagcgagcgc 19

Db 1 gcgcgtgaagcgagcgc 18

RESULT 8

AAF95308

ID AAF95308 standard; DNA; 21 BP.

XX

AC AAF95308;

XX

DT 06-JUN-2001 (first entry)

XX

DE Human gene single nucleotide polymorphism #69.

XX

KW Human; variant thrombospondin 1; variant thrombospondin 4; SNP;

KW polymorphism; vascular disease; coronary artery disease; forensics;

KW myocardial infarction; atherosclerosis; stroke; venous thromboembolism;

KW pulmonary embolism; paternity test; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Variation replace(11,A)

FT /*tag= a

FT /standard_name= "single nucleotide polymorphism"

XX

PN WO200118250-A2.

XX

XX 15-MAR-2001.

XX

PF 07-SEP-2000; 2000WO-US24503.

XX

PR 10-SEP-1999; 99US-0153357.

XX

PR 26-JUL-2000; 2000US-0220947.

XX

PR 16-AUG-2000; 2000US-0225724.

XX

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

PA

PA (MILL-) MILLENNIUM PHARM INC.

XX

XX Lander ES, Gargill M, Ireland JS, Bolk S, Daley GQ, McCarthy JJ;

XX

XX WPI; 2001-226749/23.

DR

XX Nucleic acids comprising single nucleotide polymorphisms, useful in

XX applications such as forensics, paternity testing, medicine, genetic

XX analysis and phenotype correlations to diseases such as diabetes and

XX atherosclerosis

XX

XX Examples; Page 51; 242pp; English.

PS

XX The present invention provides a method of diagnosing a vascular disease

XX in an individual, involving determining the sequence at various

XX polymorphic sites within the human thrombospondin 1 and thrombospondin 4

XX genes. The sequences at a number of polymorphic sites are also provided

XX in the specification. In particular, the method can be used in the

XX diagnosis of atherosclerosis, myocardial infarction, coronary heart

XX disease, stroke, peripheral vascular diseases, venous thromboembolism

XX and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also

XX useful in forensics, paternity testing, genetic analysis and phenotype

XX correlations to diseases. The present sequence is an example of one of

XX the human gene SNPs shown in the specification.

XX

SQ Sequence 21 BP; 7 A; 5 C; 5 G; 4 T; 0 other;

Query Match 64.0%; Score 12.8; DB 22; Length 21;

Best Local Similarity 87.5%; Pred. No. 1.9e+03;

Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ccgtgaagcggaagcga 20

Db 2 ccgtgaagcggaatgaa 17

RESULT 9

AAZ43694

ID AAZ43694 standard; DNA; 33 BP.

XX

AC AAZ43694;

XX

DT 23-FEB-2000 (first entry)

XX

DE Mass spectrometric mutation analysis primer 4.

XX

KW Primer; mass-spectrometry; genetic mutation; amplification; ss.

XX

OS Synthetic.

XX

PN DEL9824280-Al.

XX

XX 02-DEC-1999.

XX

XX 29-MAY-1998; 98DE-1024280.

XX

XX 29-MAY-1998; 98DE-1024280.

XX

XX (BRUK-) BRUKER DALTONIK GMBH.

XX

XX WPI; 2000-073581/07.

XX

XX Mass-spectrometric analysis of known gene mutations

XX

XX Example; Page 8; 16pp; German.

XX

CC This invention describes a method for mass-spectrometric analysis of

CC known genetic mutations, using modified nucleoside triphosphates to

CC improve the performance. The method comprises: (1) amplifying a DNA

CC sequence by polymerase chain reaction (PCR) using primers selected to

CC amplify a sequence containing the mutation; (2) adding a particular set

CC of modified nucleoside triphosphates (NTPs) to effect limited extension

CC of already present or newly added primers, where: (a) the extension

CC reaction stops at the next occurrence of a particular base in the DNA

CC strand being copied; (b) the extension reaction proceeds up to or past

CC the mutation site, so that wild-type amplification products will have a

CC different molecular weight from mutant amplification products; and

CC (c) the modification of the NTPs results in stabilization of the DNA

CC chains during ionization, a reduction in ion adduct formation, an

CC increase in ionization yields and/or a change in the mass of the DNA

CC chains; (3) performing the limited primer extension using an enzyme that

CC generates the complement of the DNA strand being copied; (4) performing

CC at least partial primer degradation and optionally further modification

CC of the amplification products; and (5) determining the mass of the

CC modified amplification products by mass spectrometry and assigning the

CC masses to wild type or mutant. AAZ43691-243717 represent primers used in

CC the method of the invention.

XX

SQ Sequence 33 BP; 6 A; 11 C; 11 G; 5 T; 0 other;

Query Match 63.0%;

Best Local Similarity 78.9%; Score 12.6; DB 21; Length 33;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggcgcgcgtgaagcgagcgc 19

Db 10 ggcgcgcgtgaagcgagcgc 28

```

XX PN WO200056929-A2.
XX PD 28-SEP-2000.
XX PF 17-MAR-2000; 2000WO-US07133.
XX PR 19-MAR-1999; 99US-0125251.
XX PA (CORR ) CORNELL RES FOUND INC.
XX PA (LOU ) UNIV LOUISIANA STATE.
XX PA (PURD ) PURDUE RES FOUND.
XX PI Barany F, Day JP, Hammer RP, Bergstrom DE;
XX WI: 2000-638269/61.
XX PT Coupled polymerase chain reaction-restriction endonuclease
XX PT digestion-ligase detection reaction to identify low abundance sequences
XX PT differing by single-base changes, insertion or deletion from high
XX PT abundance sequence in target sequences -
XX PS Example 7; Fig 6A; 103pp; English.
XX CC The specification describes the use of coupled polymerase chain reaction
XX CC (PCR), restriction endonuclease digestion (RED) and ligase detection
XX CC reaction (LDR) to identify one or more low abundance sequences differing
XX CC by one or more single base changes, inserts or deletions, from a high
XX CC abundance sequence, in several target nucleotide sequences. The method
XX CC involves 3 PCR reaction phases, a RED phase and an LDR phase. The method
XX CC is used to identify one or more low abundance sequences. The method is
XX CC also useful for detecting a wide variety of infectious diseases caused
XX CC by bacterial, viral, parasite and fungal infectious agent. Cancers can
XX CC also be detected by this method. The method is also used for detection,
XX CC identification and monitoring of pathogenic and indigenous municipal
XX CC waste water purification system and water reservoirs or in polluted
XX CC areas undergoing bioremediation and to detect plasmids containing genes
XX CC that can metabolise xenobiotics, to monitor specific target
XX CC microorganisms in population dynamic studies, or either to detect,
XX CC identify, or monitor specific target microorganisms modified
XX CC microorganisms in the environment and in industrial plants. The
XX CC present sequence represents a synthetic duplex DNA template, used in
XX CC the course of the invention.
XX SQ Sequence 50 BP; 11 A; 18 C; 12 G; 9 T; 0 other;

Query Match 63.0%; Score 12.6; DB 21; Length 50;
Best Local Similarity 78.9%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggccgcgtgaagcgaagc 19
   ||||| ||||| |||||
Db 12 ggccgcgtgaagcgaagc 30

RESULT 12
AAV95266/c
ID AAV95266 standard; RNA; 17 BP.
XX AC AAV95266;
XX DT 24-FEB-1999 (first entry)
XX DE Human c-fos target sequence nucleotide position 97.
XX KW Human; c-fos; hammerhead ribozyme; hairpin ribozyme; target site;
XX KW cancer; oncogene; leukaemia; neuroblastoma; diagnosis; genetic drift;
XX KW mutation; diseased cell; ss.
XX OS Homo sapiens.
XX PN WO9832846-A2.

RESULT 10
AAV36074/c
ID AAV36074 standard; DNA; 36 BP.
XX AC AAV36074;
XX DT 02-SEP-1998 (first entry)
XX DE Oligonucleotide SCE3-Stu of the specification.
XX KW Regulatory sequence; cellulase cbh1 gene; mass production;
XX KW Humicola insolens; endo-glucanase NCE4; ss.
XX OS Synthetic.
XX PN WO9811239-A1.
XX PD 19-MAR-1998.
XX PF 16-SEP-1997; 97WO-JP03268.
XX PR 13-SEP-1996; 96JP-0243695.
XX PA (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX PI Aoyagi K, Moriya T, Murakami T, Sumida N, Watanabe M;
XX WI: 1998-250959/22.
XX PT Regulatory sequence for Trichoderma viride derived cellulase cbh1
XX PT gene - for producing Humicola insolens derived endo-glucanase
XX PS Example 11; Page 29; 92pp; Japanese.
XX CC Oligonucleotides AAV36074-75 are used in the course of the invention.
XX CC The specification describes a new regulatory sequence for Trichoderma
XX CC viride derived cellulase cbh1 gene and the establishment of a system for
XX CC mass producing cellulase in moulds such as T. viride. As the regulatory
XX CC sequence of cbh1 genes originating in T. viride can highly express
XX CC objective proteins, proteins such as cellulase can be expressed. An
XX CC expression vector containing the regulatory sequence and Humicola
XX CC insolens derived endo-glucanase NCE4 DNA was produced, and used to
XX CC produce endo-glucanase at 15 grams per litre.
XX SQ Sequence 36 BP; 4 A; 11 C; 12 G; 9 T; 0 other;

Query Match 63.0%; Score 12.6; DB 19; Length 36;
Best Local Similarity 78.9%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ggccgcgtgaagcgaagc 19
   ||||| ||||| |||||
Db 24 GGAGCCATGATCGCAGGC 6

RESULT 11
AAA96608
ID AAA96608 standard; DNA; 50 BP.
XX AC AAA96608;
XX DT 08-FEB-2001 (first entry)
XX DE Nucleotide sequence of a synthetic duplex DNA.
XX KW Coupled polymerase chain reaction; PCR; ligase detection reaction; LDR;
XX KW restriction endonuclease digestion; RED; infectious disease; cancer;
XX KW waste water purification; ds.
XX OS Synthetic.

```

XX 30-JUL-1998.
PD
XX
PF 20-JAN-1998; 98WO-US01017.
XX
PR 23-JAN-1997; 97US-0037658.
XX
PA (RIBO-) RIBOZYME PHARM INC.
XX
PI Jarvis T, McSwiggen JA, Stinchcomb DF;
XX WPI; 1998-427942/36.
XX
DR Enzymatic nucleic acid molecules which specifically cleave RNA
PT derived from a c-fos gene - useful for treating conditions related
PT to levels of c-fos, especially cancer
XX
PS Claim 2; Page 50; 72pp; English.
XX
CC The present invention describes an enzymatic nucleic acid molecule which
CC specifically cleaves RNA derived from a c-fos gene. AAV95401 to AAV95540
CC and AAV95541 to AAV95584 represent hammerhead ribozymes and hairpin
CC ribozymes, respectively, which specifically cleave human c-fos. AAV95261
CC to AAV95400 and AAV95385 to AAV95628 represent human c-fos target
CC sequences. The enzymatic nucleic acid molecules can be used for treating
CC cancer associated with elevated levels of c-fos oncogene, especially
CC leukaemias, neuroblastomas and lung, breast and colon cancers. The
CC ribozymes may also be used as diagnostic tools to examine genetic drift
CC and mutations within diseased cells, or, to detect the presence of c-fos
CC RNA in a cell.
XX
SQ Sequence 17 BP; 2 A; 8 C; 3 G; 4 U; 0 other;

Query Match 61.0%; Score 12.2; DB 19; Length 17;
Best Local Similarity 82.4%; Pred. No. 3.6e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ggcgcctgaagcgaag 17
Db 17 GGCCTGTGAAGCAGAG 1

RESULT 13
AA84390
ID AA84390 standard; DNA; 19 BP.
XX
AC AA84390;
XX
DT 04-DEC-2000 (first entry)
XX
DE Cyclin D3 ribozyme binding site #1.
XX
KW Ribozyme; hairpin; hammerhead; gene therapy; vasotropic;
KW restenosis; ss.
XX
OS Mammalia.
XX
XX
XX W0200032765-A2.
PN
XX
PD 08-JUN-2000.
XX
XX
XX 06-DEC-1999; 99WO-US28772.
PF
XX
PR 04-DEC-1998; 98US-0110954.
XX
XX (IMMU-) IMMUSOL INC.
PA
XX
XX Tritz R, Welch PJ, Barber JR, Robbins JM;
PI
XX WPI; 2000-412314/35.
DR
XX
XX New hairpin and hammerhead ribozyme for inhibiting restenosis, cleaves

PT RNA encoding a cyclin or cell-cycle dependent kinase other than CDK1,
PT PCNA and Cyclin B1 -
XX
PS Disclosure; Page 76; 109pp; English.
XX
CC The present invention relates to a hairpin or hammerhead ribozyme,
CC designed to cleave RNA encoding a cyclin or cell-cycle dependent kinase
CC other than cell-cycle dependent kinases CDK1, PCNA and Cyclin B1.
CC Representative examples of ribozyme recognition sites are given in
CC AA82415 to AA86787. The ribozyme of the invention is useful for
CC inhibiting restenosis by introduction of the ribozyme into cells.
CC The ribozyme is resistant to endonuclease activity and hence is
CC efficient in restenosis treatment.
XX
SQ Sequence 19 BP; 3 A; 4 C; 7 G; 5 T; 0 other;

Query Match 61.0%; Score 12.2; DB 21; Length 19;
Best Local Similarity 82.4%; Pred. No. 3.6e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gccgtgaagcgaagca 20
Db 3 gctgtgttcgaagca 19

RESULT 14
AAF74126/C
ID AAF74126 standard; DNA; 20 BP.
XX
AC AAF74126;
XX
XX 30-APR-2001 (first entry)
XX
DE Primer #60.
XX
KW Solute carrier family 6 neurotransmitter transporter; seotonin 4;
KW SLC6A4; genotyping; allele specific oligonucleotide; ss.
XX
OS Homo sapiens.
XX
XX W0200109161-A1.
PN
XX 08-FEB-2001.
PD
XX 31-JUL-2000; 2000WO-US20638.
PF
XX 29-JUL-1999; 99US-0146290.
PR
XX (GENA-) GENAISSANCE PHARM INC.
PA
XX Denton RR, Duda A, Nandabalan K, Sanchis A, Stephens JC;
PI
XX WPI; 2001-123317/13.
DR
XX
XX New isolated polynucleotide comprising a polymorphic variant for the
PT solute carrier family 6 neurotransmitter transporter, serotonin member
PT 4 gene for identifying drugs for treating disorders related to
PT expression of the protein -
XX
XX Example 1; Page 37; 152pp; English.
XX
XX The present invention relates to a polymorphic variant of a reference
CC sequence for the solute carrier family 6 neurotransmitter
CC transporter, serotonin member 4 (SLC6A4) gene or a fragment of it
CC or a sequence complementary to the first sequence.
CC The invention is used in producing a recombinant organism
CC that can be used to express SLC6A4 for protein structure analysis and
CC binding studies. A composition comprising a genotyping oligonucleotide
CC is used to detect a polymorphism in the SLC6A4 gene.
XX
SQ Sequence 20 BP; 2 A; 8 C; 3 G; 7 T; 0 other;

Query Match 61.0%; Score 12.2; DB 22; Length 20;
Best Local Similarity 82.4%; Pred. No. 3.6e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 gccgtgaagcggaaggca 20
||| ||||| |||||
DB 19 GCGGTGAAATGAAGGCA 3

RESULT 15

AAF69762
ID AAF69762 standard; DNA; 22 BP.

XX
AC AAF69762;

XX
DT 18-APR-2001 (first entry)

XX
DE Human IL4Ralpha gene PCR primer #98.

XX
KW Polymorphism; human; interleukin 4 receptor-alpha; IL4R-alpha;
allergic disease; PCR primer; ss.

XX
OS Homo sapiens.

XX
PN WO200104270-A1.

XX
PD 18-JAN-2001.

XX
PF 13-JUL-2000; 2000WO-US19094.

XX
PR 13-JUL-1999; 99US-0143435.

XX
PA (GENA-) GENAISSANCE PHARM INC.

XX
PI Chew A, Denton RR, Duda A, Nandabalan K, Stephens JC;

XX
PI Windemuth AK;

XX
DR WPI; 2001-103078/11.

XX
PT New isolated polynucleotide useful for the identification of
therapeutics in allergic diseases is new -

XX
PS Example 1; Page 65; 188pp; English.

XX
CC The present invention relates to polymorphisms of the human interleukin 4
receptor-alpha gene (IL4R-alpha; see AAF57718 for the reference
sequence). Polynucleotides comprising polymorphic gene variants are
useful for therapeutic purposes. For example, where a patient may benefit
from expression of a particular IL4Ralpha protein isoform, an expression
vector encoding the isoform may be administered to the patient. It may
be desirable to decrease or block expression of a particular IL4Ralpha
isoform, which may be done by turning off by transforming a targeted
organ, tissue or cell population with an expression vector that expresses
high levels of untranslatable mRNA for the isoform. Specific therapeutics
identified by these methods may be useful for allergic diseases. The
present sequence is a PCR primer for human IL4R-alpha.

XX
SQ Sequence 22 BP; 9 A; 2 C; 10 G; 1 T; 0 other;

Query Match 61.0%; Score 12.2; DB 22; Length 22;
Best Local Similarity 82.4%; Pred. No. 3.7e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 gccgtgaagcggaaggca 20
||| ||||| |||||
Db 5 gcagggaagagaaggca 21

Search completed: October 2, 2001, 16:18:35
Job time: 15479 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 12:00:36 ; Search time 876.95 Seconds
(without alignments)
14.320 Million cell updates/sec

Title: US-09-757-100B-3

Perfect score: 20

Sequence: 1 ccgcgggtcacagtgtcgc 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SIDSB/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSB/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSB/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSB/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSB/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSB/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSB/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSB/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSB/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSB/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSB/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSB/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSB/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSB/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSB/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSB/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSB/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSB/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSB/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSB/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSB/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSB/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	22	Human focal adhesi
2	15	75.0	15	22	Human focal adhesi
3	13.8	69.0	31	22	Oligonucleotide sf
4	13.6	68.0	41	21	PCR primer used to
5	13.2	66.0	29	21	polymorphic fragme
6	13.2	66.0	31	19	3' PCR primer used
7	13.2	66.0	39	21	Arabidopsis thalia
8	13.2	66.0	50	21	PCR primer for DNA
9	12.8	64.0	39	11	Oligonucleotide ca
10	12.8	64.0	44	21	HIV-1 gp41 C-helic
11	12.6	63.0	29	19	Primer for ribonuc

c	12	12.6	63.0	32	21	AAZ35374	Sense primer for m
	13	12.6	63.0	47	20	AAZ01125	Probe for human PG
	14	12.2	61.0	18	21	AAA38381	Human Bts-2 phosph
c	15	12.2	61.0	29	21	AAA29338	Primer V-beta-3 fo
	16	12.2	61.0	30	20	AAZ10873	PCR primer for the
	17	12.2	61.0	31	18	AAZ63364	Primer murf2. Syn
	18	12.2	61.0	31	20	AAZ86099	PCR primer used to
	19	12.2	61.0	39	18	AAZ63369	Primer murf4. Syn
	20	12.2	61.0	39	20	AAZ86104	PCR primer used to
	21	12.2	61.0	40	21	AAZ55715	Anabaena variabill
	22	12.2	61.0	41	21	AAZ61918	PCR primer for DNA
	23	12.2	61.0	42	21	AAA37487	Arabidopsis thalia
	24	12.2	61.0	43	21	AAA37503	Yeast acyltransfer
	25	12.2	61.0	44	21	AAA37489	Arabidopsis thalia
	26	12.2	61.0	44	21	AAA37507	Yeast acyltransfer
	27	12.2	61.0	45	21	AAZ61920	PCR primer for DNA
	28	12.2	61.0	50	21	AAZ47143	Liposome membrane
c	29	12	60.0	41	19	AAV35598	Chicken IgG light
	30	12	60.0	42	21	AAA37485	Arabidopsis thalia
	31	12	60.0	43	18	AAZ66939	Cels dockerin gene
	32	12	60.0	46	21	AAA37481	Arabidopsis thalia
c	33	11.8	59.0	16	18	AAZ93016	M. tuberculosis 16
c	34	11.8	59.0	20	19	AAV45725	Human NKCC2 gene e
	35	11.8	59.0	24	20	AAZ36171	PCR primer used to
	36	11.8	59.0	29	18	AAZ92996	M. tuberculosis 16
	37	11.8	59.0	30	18	AAZ92997	M. tuberculosis 16
	38	11.8	59.0	30	18	AAZ92998	M. tuberculosis 16
	39	11.8	59.0	35	17	AAZ10707	Mycobacterium genu
	40	11.8	59.0	36	14	AAQ38115	Mycobacterium 16S
c	41	11.8	59.0	40	20	AAZ99220	M. tuberculosis 16
c	42	11.8	59.0	40	20	AAZ99221	M. bovis 16S rRNA
c	43	11.8	59.0	40	20	AAZ99225	M. smegmatis 16S r
c	44	11.8	59.0	40	20	AAZ99227	M. intracellulare
c	45	11.8	59.0	44	17	AAZ06805	Mycobacteria 16S r

ALIGNMENTS

RESULT 1
AAC65535
ID AAC65535 standard; DNA; 20 BP.
XX
AC AAC65535;
XX
DT 12-FEB-2001 (first entry)
XX
Human focal adhesion kinase antisense sequence #1.
DE
DE Human; focal adhesion kinase; FAK; signal transduction; cancer;
KW embryonic development disorder; angiogenic disorder; wound healing;
KW antisense; phosphorothioate; ss.
XX
XX Homo sapiens.
XX
XX US6133031-A.
XX
XX PD 17-OCT-2000.
XX
XX PF 19-AUG-1999; 99US-0377310.
XX
XX PR 19-AUG-1999; 99US-0377310.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Monia BP, Gaarde WA;
XX
XX WPI; 2001-006141/01.
XX
XX New antisense compounds for inhibiting focal adhesion kinase
PT expression, especially useful for inhibiting retinal
PT neovascularization, or for diagnosing and treating e.g. colon cancer -

PS Claim 3; Column 23; 30pp; English.

XX The present invention describes a number of phosphothioate antisense
CC sequences to the human focal adhesion kinase (FAK) protein. This protein
CC is involved in integrin-mediated signal transduction, and is implicated
CC in cancer, particularly colon, breast and oral tumours, embryonic
CC development disorders, angiogenic disorders and wound healing. The
CC antisense sequences, including the one shown here, can be used in the
CC treatment of all of these.

XX Sequence 20 BP; 2 A; 7 C; 8 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ccgcgggctcacagtgcg 20
|||||

Db 1 ccgcgggctcacagtgcg 20

RESULT 2

AAC65555

ID AAC65555 standard; DNA; 15 BP.

XX

AC AAC65555;

XX

DT 12-FEB-2001 (first entry)

XX

DE Human focal adhesion kinase antisense sequence #21.

XX

KW Human; focal adhesion kinase; FAK; signal transduction; cancer;

KW embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphothioate; ss.

XX

OS Homo sapiens.

XX

PN US6133031-A.

XX

PD 17-OCT-2000.

XX

PF 19-AUG-1999; 99US-0377310.

XX

PR 19-AUG-1999; 99US-0377310.

XX

PA (ISIS-) ISIS PHARM INC.

XX

PI Monia BP, Gaarde WA;

XX

DR WPI; 2001-006141/01.

XX

PT New antisense compounds for inhibiting focal adhesion kinase

PT expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

XX

PS Claim 3; Column 25; 30pp; English.

XX

XX The present invention describes a number of phosphothioate antisense
CC sequences to the human focal adhesion kinase (FAK) protein. This protein
CC is involved in integrin-mediated signal transduction, and is implicated
CC in cancer, particularly colon, breast and oral tumours, embryonic
CC development disorders, angiogenic disorders and wound healing. The
CC antisense sequences, including the one shown here, can be used in the
CC treatment of all of these.

XX

XX Sequence 15 BP; 2 A; 4 C; 7 G; 2 T; 0 other;

Query Match

Best Local Similarity 75.0%; Score 15; DB 22; Length 15;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 gcgggctcacagtgg 17
|||||

Db 1 gcgggctcacagtgg 15

RESULT 3

AAC83406/c

ID AAC83406 standard; DNA; 31 BP.

XX

AC AAC83406;

XX

DT 26-FEB-2001 (first entry)

XX

DE Oligonucleotide sfa22.

XX

KW Translational reporter vector; renilla luciferase;

KW translational recoding; ss.

XX

OS Synthetic.

XX

PN US6143502-A.

XX

PD 07-NOV-2000.

XX

PF 31-MAR-1999; 99US-0282996.

XX

PR 31-MAR-1999; 99US-0282996.

XX

PA (UTAH) UNIV UTAH RES FOUND.

XX

PI Grentzmann G, Atkins JF, Gesteland RF;

XX

DR WPI; 2001-006431/01.

XX

PT Translational reporter vector for quantification of translation
PT recoding in vivo and in vitro, comprises renilla luciferase gene,
PT polylinker for insertion of selected DNA and an out-of-frame firefly
PT luciferase gene -

XX

PS Example 4; Column 33; 21pp; English.

XX

CC The present invention relates to a translational reporter vector
CC selected from p2luc and p2luc1. The vectors have a polylinker
CC interposed between a renilla luciferase gene and a firefly luciferase
CC gene, which are out-of-frame with respect to each other but are
CC co-expressed upon recoding. The translational reporter vector is
CC useful for assaying translational recoding in vitro or in vivo.

XX

SQ Sequence 31 BP; 4 A; 10 C; 9 G; 8 T; 0 other;

Query Match

Best Local Similarity 69.0%; Score 13.8; DB 22; Length 31;

Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ccgcgggctcacagtgg 17
|||||

Db 25 CAGCGGGATCACAGTGG 9

RESULT 4

AAZ45393

ID AAZ45393 standard; DNA; 41 BP.

XX

AC AAZ45393;

XX

DT 27-MAR-2000 (first entry)

XX

XX PCR primer used to amplify the rat ACAT-like gene.

KW Acyl-CoA:cholesterol acyltransferase; ACAT; ACAT-like protein; sterol;

KW ester; triacylglycerol; fatty acyl-CoA; lipid composition; plant cell;

KW diacylglycerol acyltransferase; DAGAT; triglyceride; cancer; diabetes;

cardiopulmonary disease; heart failure; atherosclerosis; adipocytosis; leukemia; skin carcinoma; fibroblastoma; metabolic disorder; obesity; abnormal lipid metabolism; abnormal fat absorption; lipoprotein secretion; adipogenesis; PCR primer; ss.

Synthetic.
Rattus sp.

W09963096-A2.

09-DEC-1999.

04-JUN-1999; 99WO-US12541.

05-JUN-1998; 98US-0088143.

12-NOV-1998; 98US-0108389.

(CALJ) CALGENE LLC.

Lassner MW, Ruezinsky DM;

WPI; 2000-105701/09.

Novel polynucleotides used for modifying plant oil composition and for developing products for treating e.g. cancer, diabetes, cardiopulmonary disease or metabolic disorders

Example 2; Page 30; 89pp; English.

PCR primers AA245392-93 were used to amplify rat acyl-CoA:cholesterol acyltransferase (ACAT)-like gene. The ACAT-like protein is active in the formation of a sterol, ester and/or triacylglycerol from a fatty acyl-CoA and sterol and/or diacylglycerol substrate. The DNA can be used for modifying the lipid composition of plant cells. The ACAT-like protein has diacylglycerol acyltransferase (DAGAT) activity, and so the synthesis of triglycerides can be suppressed or increased using the DNA. The protein can be used to produce plant oils with a modified triglyceride content. The products can also be used to identify antagonists and agonists of DAGAT activity. Such agonists and antagonists are particularly useful in treating or ameliorating diseases associated with DAGAT activity, including diseases associated with altered cellular diacylglycerol concentration or PKC activity, including cancer, diabetes, cardiopulmonary diseases e.g. heart failure, atherosclerosis, adipocytosis, leukaemia, skin carcinoma, fibroblastoma, metabolic disorders, obesity, diseases associated with abnormal lipid metabolism, and diseases associated with abnormal fat absorption, lipoprotein secretion and adipogenesis.

Sequence 41 BP; 7 A; 13 C; 19 G; 2 T; 0 other;

Query Match 68.0%; Score 13.6; DB 21; Length 41;
Best Local Similarity 80.0%; Pred. No. 7.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccgcgggctcacagtgtctg 20
||||| | ||||| |||
Db 5 ccgcggcgccacagtggcg 24

RESULT 5

AAA04346

ID AAA04346 standard; DNA; 29 BP.

XX

AC AAA04346;

XX 22-MAY-2000 (first entry)

DT 22-MAY-2000 (first entry)

XX Polymorphic fragment of hypertension associated gene GSY1.

DE

XX

XX

KW Polymorphism; hypertension; agammaglobulinemia; diabetes insipidus;

KW Lesch-Nyhan syndrome; muscular dystrophy; Wiskott-Aldrich syndrome;

KW Fabry's disease; familial hypercholesterolemia; hereditary spherocytosis;
KW polycystic kidney disease; von Willebrand's disease; forensic;
KW tuberous sclerosis; hereditary hemorrhagica telangiectasia;
KW familial colonic polyposis; osteogenesis imperfecta; porphyria;
KW Ehlers-Danlos syndrome; ss.

XX Homo sapiens.

XX EP955382-A2.

XX 10-NOV-1999.

XX 07-MAY-1999; 99EP-0250150.

XX 07-MAY-1998; 98US-0084641.

XX 03-MAY-1999; 99US-0304232.

XX (AFFY-) AFFYMETRIX INC.

XX (UYCA-) UNIV CASE WESTERN RESERVE.

XX Fan JB, Chakravarti A, Haluska MK;

XX WPI; 2000-107928/10.

XX Novel nucleic acids containing polymorphisms used in the diagnosis of

XX hypertension

XX Claim 1; Page 33; 53pp; English.

XX The invention provides polymorphic fragments of genes associated with

XX hypertension. The nucleic acids including the polymorphic sites can be

XX used as probes or primers for expressing variant proteins. Detection of

XX the polymorphisms is useful in designing prophylactic and therapeutic

XX regimens customized to underlying abnormalities. The polymorphisms can be

XX used for association studies for hypertension, and in hypertension

XX diagnostic assays. Where the polymorphisms have strong correlation with

XX hypertension, within a gene, they are likely to have a causative role in

XX hypertension. This information can be used to find the precise role of a

XX polymorphism in the disease, and this can be used to identify potential

XX drugs which combat the disease. The polymorphisms can be tested for

XX association with other diseases e.g. agammaglobulinemia, diabetes

XX insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich

XX syndrome, Fabry's disease, familial hypercholesterolemia, polycystic

XX kidney disease, hereditary spherocytosis, von Willebrand's disease,

XX tuberous sclerosis, hereditary hemorrhagica telangiectasia, familial

XX colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, and

XX acute intermittent porphyria. The polymorphic forms can also be used in

XX forensics to identify individuals.

XX

SQ Sequence 29 BP; 6 A; 5 C; 13 G; 4 T; 1 other;

Query Match 66.0%; Score 13.2; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 1.1e+03;
Matches 15; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ccgcgggctcacagtgtctg 20
| ||| || || || || ||
Db 10 cagcgrcagacagtgttg 29

RESULT 6

AAV42721

ID AAV42721 standard; DNA; 31 BP.

XX

AC AAV42721;

XX 14-OCT-1998 (first entry)

DT 14-OCT-1998 (first entry)

XX 3' PCR primer used to amplify human furin cDNA.

DE

XX

KW Retroviral vector; gene delivery vehicle; expression; PCR primer;

KW non-immunogenic selectable marker; gene therapy; activation;

KW human; furin; ablation therapy; ss.
 XX Synthetic.
 OS Homo sapiens.
 XX WO9830709-A2.
 PN 16-JUL-1998.
 PD 14-JAN-1998; 98WO-US00715.
 XX 13-JAN-1998; 98US-0038339.
 PR 14-JAN-1997; 97US-0035473.
 XX 27-FEB-1997; 97US-0038339.
 XX (CHAD//) CHADA S.
 PA (JOLLY//) JOLLY D J.
 PA (MOOR//) MOORE M D.
 XX Chada S, Jolly DJ, Moore MD;
 PI WPI; 1998-399153/34.
 XX Non-immunogenic pro-drug activating enzyme(s) and selectable
 PT marker(s) - are used in gene therapy for the treatment of a wide
 PT variety of disorder(s)
 XX Example 8; Page 53; 121pp; English.
 PS PCR primers AAV42720-21 were used to amplify human furin cDNA (see
 CC AAV42731). The amplified product is used in the retroviral vector of the
 CC invention to encode a cell bound prodrug convertase for ablation therapy.
 CC The specification describes a gene delivery vehicle which directs
 CC expression of a non-immunogenic selectable marker or molecule which is
 CC capable of activating a previously inactive compound. Vectors expressing
 CC the markers and a heterologous sequence are useful in gene therapy. The
 CC vectors can be used to deliver a molecule into a target area where it may
 CC cause the activation of a previously inactive substance.
 XX Sequence 31 BP; 5 A; 12 C; 10 G; 4 T; 0 other;
 SQ

Query Match 66.0%; Score 13.2; DB 19; Length 31;
 Best Local Similarity 83.3%; Pred. No. 1.1e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 3 gcgggctcacagtgtcg 20
 | ||||| |||||
 Db 13 gtgggctcacagagggcg 30

RESULT 7
 AAA37483
 ID AAA37483 standard; DNA; 39 BP.
 AC AAA37483;
 XX 15-AUG-2000 (first entry)
 DT Arabidopsis thaliana acyltransferase ATAT7 PCR primer, SEQ ID NO:173.
 DE Acyltransferase; lipid synthesis; recombinant expression;
 XX membrane fluidity; cold resistance; transgenic plant;
 KW baculovirus expression; PCR primer; ss.
 XX Arabidopsis thaliana.
 OS WO200018889-A2.
 PN 06-APR-2000.
 PD 24-SEP-1999; 99WO-US22231.
 PF Savidge B, Lassner MW, Weiss JD, Post-Beittenmiller D;

PR 25-SEP-1998; 98US-0101939.
 XX (CALJ) CALGENE LLC.
 XX Lassner MW, Emig RA, Ruezinsky DM, Van Eenennaam A;
 PI WPI; 2000-303447/26.
 DR Novel acyltransferase related proteins useful for altering membrane
 XX fluidity in plant cells e.g. to induce chill tolerance -
 PT Example 7; Page 31; 126pp; English.
 PS The invention relates to nucleic acids encoding novel plant
 XX acyltransferase-like proteins (AAA37343-A37445) which comprise one of 8
 CC conserved acyltransferase motifs (AA199474-Y99481). Acyltransferases
 CC catalyse the transfer of acyl groups from a donor to a variety of
 CC substrates such as glycerides, sterols, stanols and phosphatides.
 CC Such enzymes play a key role in lipid synthesis, and thereby affect the
 CC characteristics of the plant. For example, cold-hardened plants have
 CC different lipid concentrations in the cell membrane compared to
 CC non-hardened plants, which makes the membrane more fluid and the plant
 CC more tolerant of low temperatures. The nucleic acid sequences of the
 CC invention can be used as probes or for expressing acyltransferase-like
 CC proteins in host cells e.g., for recombinant protein production. They
 CC may be expressed in plant cells to alter the lipid composition of the
 CC plant e.g., for the production of chill-resistant plants, or for altering
 CC the composition of plant oils. Sequences AAA37473-A37494 represent PCR
 CC primers used in an exemplification of the invention to amplify the
 CC Arabidopsis thaliana acyltransferase cDNAs (AAA37331-A37342) for cloning
 CC into plant and baculovirus expression constructs.
 XX Sequence 39 BP; 11 A; 9 C; 12 G; 7 T; 0 other;
 SQ

Query Match 66.0%; Score 13.2; DB 21; Length 39;
 Best Local Similarity 83.3%; Pred. No. 1.1e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 ccgcgggctcacagtgg 18
 | ||||| | |||||
 Db 5 ccgcgggccgcacaaatgg 22

RESULT 8
 AAC61924
 ID AAC61924 standard; DNA; 50 BP.
 XX AAC61924;
 AC AAC61924;
 XX 06-MAR-2001 (first entry)
 DT PCR primer for DNA encoding a prenyltransferase designated ATP78.
 DE Prenyltransferase; ATP71; ATP2; ATP3; ATP4; ATP5; ATP6; ATP7;
 XX ATP8; ATP9; ATP10; ATP11; ATP12; tocopherol; homogentisic acid;
 KW phytylpyrophosphate; 2-methyl-6phytylbenzoquinol; antioxidant;
 KW nutritional supplement; PCR primer; ss.
 XX Arabidopsis sp.
 OS WO200063391-A2.
 PN 26-OCT-2000.
 XX 14-APR-2000; 2000WO-US10368.
 PF 15-APR-1999; 99US-0129899.
 XX 30-JUL-1999; 99US-0146461.
 PR (CALJ) CALGENE LLC.
 PA Savidge B, Lassner MW, Weiss JD, Post-Beittenmiller D;

XX WPI; 2000-647519/62.
 XX An isolated nucleic acid sequence encoding prenyltransferase used to
 PT transform plant cells to increase the production of tocopherols -
 XX
 XX Example 2; Page 26; 114pp; English.
 XX PCR primers AAC61924-25 were used to amplify DNA encoding a
 CC prenyltransferase. The specification describes prenyltransferases
 CC designated ATP11, ATP2, ATP3, ATP4, ATP5, ATP6, ATP7, ATP8,
 CC ATP9, ATP10, ATP11, and ATP12. The biosynthesis of alpha-tocopherol
 CC in higher plants involves the condensation of homogentisic acid and
 CC phytylpyrophosphate to form 2-methyl-6-phytylbenzoquinol, which can form
 CC various tocopherols. The prenyltransferase polynucleotides are useful in
 CC transforming host cells to alter the expression of prenyltransferase in
 CC these cells. The transformed cells are used in the production of
 CC tocopherols which are of use in the pharmaceutical industry as
 CC antioxidants and also in the food industry as nutritional supplements.
 XX
 SQ Sequence 50 BP; 9 A; 16 C; 14 G; 11 T; 0 other;

Query Match 66.0%; Score 13.2; DB 21; Length 50;
 Best Local Similarity 83.3%; Pred. No. 1.1e+03;
 Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccgcgggctcacagtgt 18
 ||||| | |||| |
 Db 5 ccgcggcgccacaatggt 22

RESULT 9
 AAQ04938
 ID AAQ04938 standard; DNA; 39 BP.
 XX
 AC AAQ04938;
 XX
 XX 24-OCT-1990 (first entry)
 XX
 DE Oligonucleotide carrying mutation for factor VIII gene.
 XX
 KW Human factor VIII analogue; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT mutation 18..23
 FT /*tag= a
 XX
 XX WO9005530-A.
 XX
 XX 31-MAY-1990.
 XX
 XX 14-NOV-1989; 89WO-0005049.
 XX
 XX 14-NOV-1988; 88US-0270882.
 XX
 XX (GENE-) GENETICS INST INC.
 XX
 XX Kaufman RJ, Pittman DD;
 XX
 XX WPI; 1990-193265/25.
 XX
 XX New hybrid DNA encoding hybrid procoagulant proteins -
 PT prep'd. by modifying DNA encoding human factor VIII.
 XX
 XX Disclosure; ; pp; English.
 XX
 XX Factor VIII analogue is sufficiently mutated from the original gene
 CC that it is not recognised by blood Abs of the patient. The analogue
 CC is composed of human FVIII but carries the B-domain of human FV in
 CC place of the FVIII B-domain.

CC Oligonucleotides can be used to alter the profile of the blood
 CC factor without significantly altering its activity.
 XX
 SQ Sequence 39 BP; 9 A; 11 C; 11 G; 8 T; 0 other;

Query Match 64.0%; Score 12.8; DB 11; Length 39;
 Best Local Similarity 87.5%; Pred. No. 1.8e+03;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ccgcgggctcacagtgg 17
 |||| | |||| |
 Db 19 ccgcgtgctcagagtgg 34

RESULT 10
 AAA72106
 ID AAA72106 standard; DNA; 44 BP.
 XX
 AC AAA72106;
 XX
 XX 24-NOV-2000 (first entry)
 XX
 DE HIV-1 gp41 C-helical domain 5' PCR primer.
 XX
 KW HIV-1; gp41; N-helical domain; heptad repeat region; C-helical domain;
 KW gp41 transmembrane-proximal amphipathic alpha-helical segment;
 KW core 6-helix bundle; viral entry inhibition; immunogenic;
 KW antibody; humoral response; broad spectrum vaccine; anti-HIV;
 KW envelope glycoprotein; prophylaxis; therapy; PCR primer; ss.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 XX WO200040616-A1.
 XX
 XX 13-JUL-2000.
 XX
 XX 10-JAN-2000; 2000WO-US00456.
 XX
 XX 08-JAN-1999; 99US-0115404.
 XX
 XX 07-JAN-2000; 2000US-0480336.
 XX
 XX (WILD/) WILD C T.
 XX (WEISS/) WEISS C D.
 XX
 XX Wild CT, Weiss CD;
 XX
 XX WPI; 2000-465959/40.
 XX
 XX Raising neutralizing antibody response to human immunodeficiency virus,
 PT comprises administering a polypeptide capable of forming a stable
 PT coiled-coil solution structure -
 XX
 XX Example 5; Page 54; 97pp; English.
 XX
 XX The invention relates to raising a neutralising antibody response to a
 CC broad spectrum of HIV (human immunodeficiency virus) strains and
 CC isolates, comprising the administration of a peptide which corresponds
 CC to or mimics highly conserved portions of the gp41 envelope glycoprotein
 CC which are important in mediating the process of viral entry into host
 CC cells. Such peptides can correspond to or mimic the coiled coil
 CC solution structure of the N-helical domain (the heptad repeat
 CC region), or can correspond or mimic the C-helical domain (the
 CC transmembrane-proximal amphipathic alpha-helical segment), or the
 CC gp41 core 6-helix bundle, which is formed by the interaction of
 CC the N- and C-helical domains of three gp41 proteins. The peptides
 CC can be administered either singly or as a combination (particularly
 CC a combination of N-helical and C-helical peptides), and can be
 CC multimerised. For example, N- and C-helical domain peptides can be
 CC alternately linked together to form a peptide which mimics the core
 CC 6-helix bundle. Administration of the peptide(s) generates a humoral
 CC response, with the production of antibodies against gp41 structures
 CC involved in viral entry. As these portions of gp41 are well conserved,

CC such antibodies may be effective against a broad range of HIV strains
 CC and isolates. The peptide compositions may be administered as a
 CC prophylactic or therapeutic vaccine to generate antibodies which reduce
 CC or inhibit the ability of HIV to infect uninfected cells. A composition
 CC comprising polyclonal or monoclonal antibodies can be administered to
 CC reduce HIV infection of uninfected cells. Antibodies raised against
 CC entry-relevant gp41 structures may also be used therapeutically and as
 CC tools to further elucidate the mechanism of HIV cell entry.
 CC Sequences AAA72104-A72107 represent PCR primers used in an
 CC exemplification of the invention to construct DNA (AAA72102) encoding a
 CC protein construct (RAB14603) comprising the N- and C-helical domains of
 CC HIV-1 gp41 separated by a short peptide linker.
 XX
 SQ Sequence 44 BP; 13 A; 10 C; 15 G; 6 T; 0 other;

Query Match 64.0%; Score 12.8; DB 21; Length 44;
 Best Local Similarity 87.5%; Pred. No. 1.8e+03;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 cgcggggtcacagtgg 17
 ||| ||||| |||||
 Db 8 cgcggggtcacagtgg 23

RESULT 11
 AAV59089
 ID AAV59089 standard; DNA; 29 BP.
 XX
 AC AAV59089;
 DT 15-JAN-1999 (first entry)
 XX
 DE Primer for ribonucleotide reductase R1 coding sequence.
 XX
 KW Ribonucleotide reductase R1; tumourigenicity modulation; therapy;
 KW neoplastic cell; tumour growth; tumour regression; PCR primer; mouse; ss.
 XX
 OS Synthetic.
 OS Mus sp.
 XX
 PN WO9841231-AL.
 PD 24-SEP-1998.
 XX
 PF 18-MAR-1998; 98WO-CA00242.
 XX
 PR 02-OCT-1997; 97US-0060669.
 PR 19-MAR-1997; 97US-0040837.
 XX
 PA (GENE-) GENESENSE TECHNOLOGIES INC.
 XX
 PI Wright JA, Young AH;
 XX
 DR WPI; 1998-520956/44.
 XX
 PT Use of ribonucleotide reductase R - for modulating the
 PT tumourigenicity and metastatic potential of neoplastic cells in a
 PT mammal, particularly for inhibiting tumour growth
 XX
 PS Example; Page 21; 56pp; English.
 XX
 CC This sequence represents a PCR primer for the mouse ribonucleotide
 CC reductase R1 (RRR1) gene. The gene and the protein it encodes can be
 CC used in the methods of the invention for modulating tumourigenicity of
 CC neoplastic cells in a mammal. The RRR1 can slow tumour growth and cause
 CC tumour regression and reduction of tumourigenicity and metastatic
 CC potential.
 XX
 SQ Sequence 29 BP; 3 A; 11 C; 7 G; 8 T; 0 other;

Query Match 63.0%; Score 12.6; DB 19; Length 29;
 Best Local Similarity 78.9%; Pred. No. 2.2e+03;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Best Local Similarity 78.9%; Pred. No. 2.2e+03;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ccgcggggtcacagtggtc 19
 ||| ||| ||||| |||||
 Db 2 cctcgcgtgacagtcgctc 20

RESULT 12
 AAZ35374/C
 ID AAZ35374 standard; DNA; 32 BP.
 XX
 AC AAZ35374;
 DT 27-MAR-2000 (first entry)
 XX
 DE Sense primer for methylthioadenosine phosphorylase gene exons 5-8.
 XX
 KW Chromosome 9p21; human; cancer; tumour; diagnosis; prognosis;
 KW methylthioadenosine phosphorylase; MTPAP; glioma; melanoma;
 KW primary lymphoid malignancy; non-small cell lung cancer;
 KW head and neck cancer; ovarian cancer; bladder cancer;
 KW chondrosarcoma; PCR primer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9967634-AL.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13981.
 XX
 PR 23-JUN-1998; 98US-0090411.
 PR 17-JUN-1999; 99US-0335231.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Carson DA, Schmid M, Carrera CJ;
 XX
 DR WPI; 2000-126650/11.
 XX
 PT Diagnosing and determining prognosis for cancer causatively associated
 PT with derangements of chromosome 9p21 -
 XX
 PS Claim 7; Page 16; 55pp; English.
 XX
 CC This sense primer is designed for the PCR amplification of exons
 CC 5-8 of the human methylthioadenosine phosphorylase (MTPAP) gene on
 CC chromosome 9p21 (see AAZ35374). The primer is used, with an
 CC antisense primer (see AAZ35375), in a PCR amplification (505 bp
 CC product) in step (a) of claimed methods for diagnosis of, and
 CC determining a prognosis for, cancer causatively associated with
 CC derangements of chromosome 9p21. Step (a) of these methods
 CC involves determining whether any portion of the 9p21 chromosome
 CC including and telomeric to STS 3.21 is deleted, or whether any
 CC portion of the gene encoding MTPAP is deleted. Step (b) involves
 CC determining whether any portion of 9p21 centromeric to STS 3.21
 CC is deleted, or whether any portion of the gene coding for p16 is
 CC deleted. The results are indicative of a cancer at an early stage
 CC or advanced stage of tumour development. Primer pairs (see
 CC AAZ35374-75) are provided for use in the claimed methods, where the
 CC cancer is a glioma, primary lymphoid malignancy, non-small cell
 CC lung cancer, melanoma, head and neck cancer, ovarian cancer,
 CC bladder cancer or a chondrosarcoma (claimed).
 XX
 SQ Sequence 32 BP; 8 A; 9 C; 6 G; 9 T; 0 other;

Query Match 63.0%; Score 12.6; DB 21; Length 32;
 Best Local Similarity 78.9%; Pred. No. 2.2e+03;
 Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ccgcggggtcacagtggtc 19


```

XX AAA29338;
XX AC
XX DT
XX DE 12-SEP-2000 (first entry)
XX DE Primer V-beta-3 for MAGE-1 specific TCR-alpha chain cloning.
XX KW human leukocyte antigen; HLA-A1; melanoma; neoplastic; MAGE-1; TCR; MHC;
XX KW soluble; major histocompatibility complex; antigen; T cell receptor;
XX KW lymphocyte; tumour; cytostatic; anti-microbial; immunosuppressive;
XX KW primer; alpha chain; ss.
XX OS
XX OS Homo sapiens.
XX PN WO200031239-A1.
XX PD 02-JUN-2000.
XX PF 18-NOV-1999; 99WO-IL00622.
XX PR 19-NOV-1998; 98IL-0127142.
XX PA (YEDA ) YEDA RES & DEV CO LTD.
XX PA (BOLH/) BOLHUIS R L H.
XX PI Bolhuis RLH, Eshhar Z, Willemssen RA;
XX WPI; 2000-451678/39.
XX PT Immune cells with predefined specificities useful for treating melanoma
XX PT and immune diseases
XX PS Example 1; Page 29; 51pp; English.
XX CC AAA29335-38 are oligonucleotide primers used to synthesize the alpha and
XX CC beta chain cDNA for a MAGE-1 specific T cell receptor from the CD8
XX CC positive cytotoxic T lymphocyte (CTL) clone 82/30. Novel immune cells
XX CC with predefined specificity, are produced by either complexing the cells
XX CC with an antigen-specific MHC-restricted TCR or transfecting the cells
XX CC with an antigen-specific MHC-restricted chimeric TCR gene. The
XX CC antigen-specific MHC-restricted TCR can be complexed with lymphocytic
XX CC cells for treatment of a tumour. Alternatively, the autologous
XX CC lymphocytes can be transfectd with an antigen-specific MHC-restricted
XX CC chimeric TCR gene encoding a single chain TCR (scFv-TCR) which binds to
XX CC an antigen associated with the tumour and a segment encoding a signal
XX CC transducing element of an immune cell. Compositions comprising the
XX CC immune cells may be used for the treatment of cancer (especially
XX CC melanomas, if the TCR binds to the MAGE-1 antigen), infectious diseases,
XX CC autolimmune disease and/or graft rejection.
XX SQ Sequence 29 BP; 4 A; 10 C; 11 G; 4 T; 0 other;

```

```

Query Match      61.0%; Score 12.2; DB 21; Length 29;
Best Local Similarity 82.4%; Pred. No. 3.4e+03;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Oy 4 cgggctcacagtgtcg 20
   | | | | | | | | | |
Db 24 CCGGCTGACAGTGTCTCG 8

```

```

Search completed: October 2, 2001, 16:18:33
Job time: 15477 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:55 ; Search time 417.38 Seconds
(without alignments)
6.804 Million cell updates/sec

Title: US-09-757-100B-33

Perfect score: 15

Sequence: 1 tggcttattctcagt 15

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_NA:*
- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
 - 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
 - 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
 - 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
 - 5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15	3	US-09-377-310-33
2	15	100.0	20	3	US-09-377-310-13
c 3	13	86.7	21	4	US-08-462-467B-28
c 4	12.4	82.7	27	1	US-08-459-568-11
c 5	12.4	82.7	27	2	US-08-399-411-11
c 6	12.4	82.7	27	3	US-08-516-859A-11
c 7	12.4	82.7	30	3	US-08-937-271-39
c 8	12	80.0	21	3	US-09-087-232A-2
9	11.8	78.7	22	4	US-09-116-498-16
10	11.8	78.7	30	1	US-07-884-811-10
11	11.8	78.7	30	1	US-07-885-971-10
12	11.8	78.7	30	1	US-08-087-783A-10
13	11.8	78.7	30	1	US-08-194-088B-10
14	11.8	78.7	30	1	US-08-435-501-9
15	11.8	78.7	30	1	US-08-435-764-9
16	11.8	78.7	30	1	US-08-792-078-9
17	11.8	78.7	30	2	US-08-194-087-10
18	11.8	78.7	30	5	PCT-US93-04648-10
19	11.8	78.7	30	5	PCT-US93-04717-9
20	11.4	76.0	21	1	US-08-271-946A-7
21	11.4	76.0	21	1	US-08-271-942A-7
22	11.4	76.0	21	3	US-08-779-916A-7
23	11.4	76.0	21	5	PCT-US95-08604-7
24	11.4	76.0	21	5	PCT-US95-08606-7
25	11.4	76.0	27	3	US-08-716-190-14
26	11.4	76.0	31	3	US-08-961-083-272
27	11.4	76.0	43	3	US-08-832-985-49

c	28	11	73.3	17	3	US-08-467-023-105	Sequence 105, Appl
	29	11	73.3	24	1	US-08-198-431-14	Sequence 14, Appl
	30	11	73.3	24	1	US-08-564-955-14	Sequence 14, Appl
	31	11	73.3	24	2	US-08-537-874-14	Sequence 14, Appl
	32	11	73.3	24	3	US-08-621-859-14	Sequence 14, Appl
	33	11	73.3	24	3	US-09-100-856A-14	Sequence 14, Appl
	34	11	73.3	24	4	US-09-075-511-14	Sequence 14, Appl
	35	11	73.3	24	4	US-09-099-015-14	Sequence 14, Appl
	36	11	73.3	31	6	5164180-1	Patent No. 5164180
c	37	11	73.3	41	1	US-08-198-431-13	Sequence 13, Appl
c	38	11	73.3	41	1	US-08-564-955-13	Sequence 13, Appl
c	39	11	73.3	41	2	US-08-537-874-13	Sequence 13, Appl
c	40	11	73.3	41	3	US-08-621-859-13	Sequence 13, Appl
c	41	11	73.3	41	3	US-09-100-856A-13	Sequence 13, Appl
c	42	11	73.3	41	4	US-09-075-511-13	Sequence 13, Appl
c	43	11	73.3	41	4	US-09-099-015-13	Sequence 13, Appl
	44	10.8	72.0	16	2	US-08-874-266-5	Sequence 5, Appl
	45	10.8	72.0	17	1	US-08-217-016-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-377-310-33
; Sequence 33, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-33

Query Match 100.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 tggcttattctcagt 15
Db 1 tggcttattctcagt 15

RESULT 2
US-09-377-310-13
; Sequence 13, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-13

Query Match 100.0%; Score 15; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tggcttatcttcagt 15
| | | | | | | | | | | | | | | |
Db 3 tggcttatcttcagt 17

RESULT 3
US-08-462-467B-28/c
; Sequence 28, Application US/08462467B
; Patent No. 6210899
; GENERAL INFORMATION:
; APPLICANT: Rosenbaum, Jan S
; TITLE OF INVENTION: The Use of a BMP Protein Receptor
; TITLE OF INVENTION: Complex for Screening Bone Metabolism Actives and Cells
; TITLE OF INVENTION: Co-transfected with a Type II BMP Receptor and a Type I
; TITLE OF INVENTION: BMP Receptor
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Procter & Gamble Company
; STREET: 11810 East Miami River Road
; CITY: Ross
; STATE: OH
; COUNTRY: USA
; ZIP: 45061
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462.467B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hersko, Bart S.
; REGISTRATION NUMBER: 32,572
; REFERENCE/DOCKET NUMBER: 5474R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 627-0633
; TELEFAX: (513) 627-0260
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-462-467B-28

Query Match 86.7%; Score 13; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 tggcttatcttca 13
| | | | | | | | | | | | | | | |
Db 18 TGGCTTATCTTCA 6

RESULT 4
US-08-459-568-11/c
; Sequence 11, Application US/08459568
; Patent No. 5811304
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi

; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459.568
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/399,411
; FILING DATE: 06-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 1264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-459-568-11

Query Match 82.7%; Score 12.4; DB 1; Length 27;
Best Local Similarity 92.9%; Pred. No. 1.3e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 tggcttatcttcag 14
| | | | | | | | | | | | | | | |
Db 27 TGGCTTTTCTTCAG 14

RESULT 5
US-08-399-411-11/c
; Sequence 11, Application US/08399411
; Patent No. 5831008
; GENERAL INFORMATION:
; APPLICANT: Huang, Shi
; TITLE OF INVENTION: Retinoblastoma Protein - Interacting
; TITLE OF INVENTION: Zinc Finger Proteins
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/399,411
; FILING DATE: 06-MAR-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1264
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-399-411-11

Query Match 82.7%; Score 12.4; DB 2; Length 27;
Best Local Similarity 92.9%; Pred. No. 1.3e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tggcttatcttcag 14
||||| |||||
Db 27 TGGCTTTCTTCAG 14

RESULT 6

US-08-516-859A-11/c
Sequence 11, Application US/08516859A
Patent No. 6069231
GENERAL INFORMATION:
APPLICANT: Huang, Shi
TITLE OF INVENTION: Retinoblastoma Protein - Interacting
TITLE OF INVENTION: Zinc Finger Proteins
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/516,859A
FILING DATE: 18-AUG-1995
CLASSIFICATION: 514
PRIOR APPLICATION NUMBER: US 08/399,411
FILING DATE: 06-MAR-1995
INFORMATION DATA:
APPLICATION NUMBER: US 08/292,683
FILING DATE: 18-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1776
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-516-859A-11

Query Match 82.7%; Score 12.4; DB 3; Length 27;
Best Local Similarity 92.9%; Pred. No. 1.3e+02;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 tggcttatcttcag 14
||||| |||||
Db 27 TGGCTTTCTTCAG 14

RESULT 7

US-08-937-271-39
Sequence 39, Application US/08937271
Patent No. 6063386
GENERAL INFORMATION:
APPLICANT: Dale, James B.
APPLICANT: Lederer, James W.
TITLE OF INVENTION: RECOMBINANT MULTIVALENT M PROTEIN
TITLE OF INVENTION: VACCINE
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/937,271
FILING DATE: 15-SEP-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Rosenman, Stephen J.
REGISTRATION NUMBER: 43,058
REFERENCE/DOCKET NUMBER: 481112.405C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-937-271-39

Query Match 82.7%; Score 12.4; DB 3; Length 30;
Best Local Similarity 92.9%; Pred. No. 1.3e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tggcttatcttcag 14
||||| |||||
Db 10 TGGCTTATCTTCG 23

RESULT 8

US-09-087-232A-2/c
Sequence 2, Application US/09087232A
Patent No. 6153431
GENERAL INFORMATION:
APPLICANT: Quillent et al.
TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR
TITLE OF INVENTION: VARIANTS ASSOCIATED WITH RESISTANCE TO VIRUS INFECTION.
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA

Query Match 82.7%; Score 12.4; DB 3; Length 27;
Best Local Similarity 92.9%; Pred. No. 1.3e+02;

ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/087.232A
FILING DATE: 28 MAY 1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/048,057
FILING DATE: 30 MAY 1997
ATTORNEY/AGENT INFORMATION:
NAME: KOLE, LISA B.
REGISTRATION NUMBER: 35,225
REFERENCE/DOCKET NUMBER: AP 31115
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 408-2628
TELEFAX: (212) 765-2519
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "PRIMER"
US-09-087-232A-2

Query Match 80.0%; Score 12; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 ggcttatcttca 13
Db 13 GGGTTATCTTCA 2
RESULT 9
US-09-116-498-16
Sequence 16, Application US/09116498
Patent No. 6251582
GENERAL INFORMATION:
APPLICANT: Littman, Dan R.
Deng, Hongkui
Unutmaz, Derya
Ramani, Vineet N.K.
TITLE OF INVENTION: NOVEL ALTERNATIVE G-COUPLED RECEPTORS
ASSOCIATED WITH RETROVIRAL ENTRY INTO CELLS, METHODS OF
IDENTIFYING THE SAME, AND DIAGNOSTIC AND THERAPEUTIC USES
THEREOF

NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/116.498
FILING DATE: 16-Jul-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotides"
HYPOTHETICAL: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-116-498-16
Query Match 78.7%; Score 11.8; DB 4; Length 22;
Best Local Similarity 86.7%; Pred. No. 2.7e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 tggcttatcttcagt 15
Db 4 TGGCTTATCATCAAT 18
RESULT 10
US-07-884-811-10
Sequence 10, Application US/07884811
Patent No. 5316921
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/884,811
FILING DATE: 19920518
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 755.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 bases
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-884-811-10

Query Match 78.7%; Score 11.8; DB 1; Length 30;
Best Local Similarity 86.7%; Pred. No. 2.8e+02;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tggcttatcttcagt 15
||| |||||
Db 7 TGACATATCTTCAGT 21

RESULT 11

US-07-885-971-10
; Sequence 10, Application US/07885971
; Patent No. 5328837
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,971
; FILING DATE: 19920518
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-885-971-10

Query Match 78.7%; Score 11.8; DB 1; Length 30;

Best Local Similarity 86.7%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tggcttatcttcagt 15
||| |||||
Db 7 TGACATATCTTCAGT 21

RESULT 12

US-08-087-783A-10
; Sequence 10, Application US/08087783A
; Patent No. 5547856
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California

; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,783A
; FILING DATE: 13-Jul-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0755779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-087-783A-10

Query Match 78.7%; Score 11.8; DB 1; Length 30;

Best Local Similarity 86.7%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 tggcttatcttcagt 15
||| |||||
Db 7 TGACATATCTTCAGT 21

RESULT 13

US-08-194-088B-10
; Sequence 10, Application US/08194088B
; Patent No. 5580963
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,088B
; FILING DATE: 09-FEB-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gallegos, R. Thomas
; REGISTRATION NUMBER: 32,692

REFERENCE/DOCKET NUMBER: 755D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-2614
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-194-088B-10

Query Match 78.7%; Score 11.8; DB 1; Length 30;
Best Local Similarity 86.7%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tggcttatcttcagt 15
||| |||||
Db 7 TGACATATCTTCAGT 21

RESULT 14
US-08-435-501-9
Sequence 9, Application US/08435501
Patent No. 5684136
GENERAL INFORMATION:
APPLICANT: Genentech, Inc., Paul J. Godowski
TITLE OF INVENTION: Receptor Activation
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,501
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/884811
FILING DATE: 18-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/885971
FILING DATE: 18-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/950572
FILING DATE: 21-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 773P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-435-501-9

Query Match 78.7%; Score 11.8; DB 1; Length 30;
Best Local Similarity 86.7%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tggcttatcttcagt 15
||| |||||
Db 7 TGACATATCTTCAGT 21

RESULT 15
US-08-435-764-9
Sequence 9, Application US/08435764
Patent No. 5763584
GENERAL INFORMATION:
APPLICANT: Genentech, Inc., Paul J. Godowski
TITLE OF INVENTION: Receptor Activation
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,764
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/884811
FILING DATE: 18-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/885971
FILING DATE: 18-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/950572
FILING DATE: 21-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Ginger R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: 773P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-435-764-9

Query Match 78.7%; Score 11.8; DB 1; Length 30;
Best Local Similarity 86.7%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 tggcttatcttcagt 15
||| |||||
Db 7 TGACATATCTTCAGT 21

Search completed: October 2, 2001, 16:03:55
Job time: 14599 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:48 ; Search time 876.95 seconds
(without alignments)
10.740 Million cell updates/sec

Title: US-09-757-100b-33

Perfect score: 15

Sequence: 1 tggcttatcttcagt 15

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15	22 AAC65565	Human focal adhesi
2	15	100.0	20	19 AAV08375	PCR primer for Cel
3	15	100.0	20	22 AAC65545	Human focal adhesi
4	13	86.7	21	17 AAT28037	PCR primer for mou
5	13	86.7	21	17 AAT27230	Antisense primer f
6	12.4	82.7	20	20 AAX94668	PCR primer used to
7	12.4	82.7	21	22 AAF95414	Human gene single
8	12.4	82.7	27	21 AAG60107	Rat RIZ PCR primer
9	12.4	82.7	30	21 AAX57918	Streptococcus pyog
10	12	80.0	19	22 AAF76114	CCR5/CCR2b PCR pri
11	12	80.0	21	20 AAV84128	HIV-1 co-receptor

c 12	11.8	78.7	21	20 AAV08563	Primer for Transge
13	11.8	78.7	22	20 AAX07298	Human SIV/HIV rece
c 14	11.8	78.7	23	21 AAA66237	Dog genomic marker
15	11.8	78.7	30	14 AAQ52624	Primer delatK3 hu
16	11.8	78.7	30	20 AAX15645	Primer used to gen
17	11.4	76.0	21	17 AAT12845	PCR 5', primer for
18	11.4	76.0	21	17 AAT11444	Retinoblastoma gen
19	11.4	76.0	27	16 AAT00441	N-terminal primer
20	11.4	76.0	31	19 AAV27482	Streptococcus pneu
21	11.4	76.0	43	20 AAV84748	Antibody heavy cha
c 22	11	73.3	17	15 AAQ55632	Juniperus Cry j i
c 23	11	73.3	20	13 AAQ29005	Sequence of PCR pr
24	11	73.3	24	18 AAT73032	Mutagenic PCR prim
25	11	73.3	24	19 AAV54903	Primer used to int
26	11	73.3	31	13 AAQ26930	Probe for the nove
27	11	73.3	31	18 AAT62463	G-alpha protein GP
28	11	73.3	31	18 AAT62485	yeast alpha-factor
29	11	73.3	31	18 AAT49593	Bacillus thuringie
c 30	11	73.3	41	18 AAT73031	Mutagenic PCR prim
c 31	11	73.3	41	19 AAV54902	Primer used to int
32	11	73.3	43	18 AAT62479	STE2 gene 3' PCR p
33	10.8	72.0	15	19 AAV00716	ATM region b (nt 5
34	10.8	72.0	17	20 AAX15040	Antisense PCR prim
35	10.8	72.0	20	13 AAQ32816	Microsatellite rep
36	10.8	72.0	20	15 AAQ57839	Primer pair 14A W
c 37	10.8	72.0	20	16 AAQ76318	Human beta-lactogl
38	10.8	72.0	20	19 AAV22500	Antisense oligonuc
c 39	10.8	72.0	20	19 AAV12309	Ribonucleotide red
c 40	10.8	72.0	20	20 AAX92612	PCR primer used to
41	10.8	72.0	20	21 AAA90730	Ribonucleotide red
c 42	10.8	72.0	20	21 AAZ88170	Mouse ssH2B PCR pr
c 43	10.8	72.0	21	20 AAX59982	Oligonucleotide pr
c 44	10.8	72.0	22	21 AAA47518	Primer for amplify
c 45	10.8	72.0	23	14 AAQ52566	HPV-16 target sequ

ALIGNMENTS

RESULT 1
AAC65565
ID AAC65565 standard; DNA; 15 BP.
XX
AC AAC65565;
XX
XX 12-FEB-2001 (first entry)
DT
XX Human focal adhesion kinase antisense sequence #31.
DE
XX Human; focal adhesion kinase; FAK; signal transduction; cancer;
KW embryonic development disorder; angiogenic disorder; wound healing;
KW antisense; phosphorothioate; ss.
XX
OS Homo sapiens.
XX
PN US6133031-A.
XX
PD 17-OCT-2000.
XX
XX 19-AUG-1999; 99US-0377310.
PF 19-AUG-1999; 99US-0377310.
XX
PR 19-AUG-1999; 99US-0377310.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Monia BP, Gaarde WA;
XX
XX WPI; 2001-006141/01.
XX
XX New antisense compounds for inhibiting focal adhesion kinase
PT expression, especially useful for inhibiting retinal
PT neovascularization, or for diagnosing and treating e.g. colon cancer -
XX

Claim 15; Column 25; 30pp; English.

xx The present invention describes a number of phosphothioate antisense
CC sequences to the human focal adhesion kinase (FAK) protein. This protein
CC is involved in integrin-mediated signal transduction, and is implicated
CC in cancer, particularly colon, breast and oral tumours, embryonic
CC development disorders, angiogenic disorders and wound healing. The
CC antisense sequences, including the one shown here, can be used in the
CC treatment of all of these.

Sequence 15 BP; 2 A; 3 C; 3 G; 7 T; 0 other;
SQ

Query Match 100.0%; Score 15; DB 22; Length 15;

Best Local Similarity	100.0%;	Pred. No. 22;	
Matches	15;	Conservative	0; Mismatches
			0; Indels
			0; Gaps

Qy 1 tggcttatcttcagt 15

RESULT 2

AAV08375/C

ID AAV08375 standard; DNA; 20 BP.

XX

AC AAV08375;

XX
DT 08-FEB-1999 (first entry)

XX

DE PCR primer

[illegible]

KY
 WY
 YY
 cellulose sy

Q5. Synthetic.

OS *Gossypium hirsutum*.

XX
XX

PN EP875575-A2.

XX 2 3 4 5 6 7 8 9 10 11 12

PD 04-NOV-1998.

31-MAR-1998;	98EP-0302489.
XX	
01-APR-1997;	97JP-0083133.
PR	

XX

XX	(N1SN) NISSHINBO IND INC.
PA	
XX	
PI	Aotsuka S, Hasegawa O, Hayashi T, Ihara Y;
XX	
DR	WPI; 1998-559440/48.
XX	
PT	New DNA encoding cotton cellulose synthase proteins - useful for
PT	controlling cellulose synthesis in prokaryotic or eukaryotic cells
XX	
PS	Example 5; Page 45; 51pp; English.
XX	
CC	This sequence represents a PCR primer for DNA encoding a cellulose
CC	synthase of the invention. The DNA can be used for controlling cellulose
CC	synthesis in a cell by introducing the DNA into the cell, expressing RNA
CC	with a nucleotide sequence homologous to the DNA or a nucleotide sequence
CC	complementary to the DNA. The cells are prokaryotic cells
CC	i.e. acetobacterium and/or eukaryotic cells such as yeasts
CC	i.e. Saccharomyces spp., plant cells (cotton plants) and cultured cells
CC	of mammals etc.

Sequence 20 BP: 10 A: 3 C: 5 G: 2 T: 0 other:
SQ

Query Match

Best Local Similarity 100.0%; Pred. No. 22;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 tggcttatcttcagt 15

OS Synthetic.
 PN WO9614579-A1.
 XX
 XX 17-MAY-1996.
 PD
 XX
 XX 30-OCT-1995; 95WO-US14027.
 PF
 XX
 XX 05-JUN-1995; 95US-0462467.
 PR
 XX 04-NOV-1994; 94US-0334178.
 XX
 XX (PROC) PROCTER & GAMBLE CO.
 PA
 XX
 XX Rosenbaum JS;
 PI
 XX WPI; 1996-251887/25.
 DR
 XX Assays for bone morphogenetic protein activities - using complex of
 PT BMP type I receptor kinase protein and BMP receptor kinase protein
 PT BRK-3
 XX
 XX Example 9; Page 32; 101pp; English.
 PS
 XX A PCR primer (AAT28037) includes the reverse complement of a
 CC partial cDNA clone isolated from NIH3T3 mouse embryonic
 CC fibroblasts. It was used with an anchor primer (AAT28038) to
 CC amplify the 5' end of mouse bone morphogenetic protein (BMP) type
 CC receptor kinase protein-3 (BRK-3) cDNA. Following further PCR
 CC amplification and cloning into pJT6, a cDNA clone (AAT28020) coding
 CC for full-length murine BRK-3 (AA95224) was obt'd. The PCR primer is
 XX the reverse complement of nucleotides 968-948 of this clone.
 XX
 SQ Sequence 21 BP; 6 A; 4 C; 6 G; 5 T; 0 other;

Query Match 86.7%; Score 13; DB 17; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggcttatcttca 13
 |||||
 Db 18 TGGCTTATCTTCA 6

RESULT 5
 AAT27230/c
 ID AAT27230 standard; DNA; 21 BP.
 AC
 XX
 XX 23-AUG-1996 (first entry)
 DT
 XX
 XX Antisense primer for mouse BRK-3 gene 5'-end amplification.
 DE
 XX
 XX Primer; antisense; 5'-RACE; PCR; polymerase chain reaction;
 KW mouse; bone morphogenetic protein receptor kinase-3; antibody;
 KW diagnostic; bone disorder; osteogenic; reporter gene;
 KW bone morphogenetic protein-agonist; drug screening;
 KW bone morphogenetic protein-antagonist; hormone-responsive element; ss.
 OS
 XX Synthetic.
 XX
 XX WO9614412-A2.
 PN
 XX
 XX 17-MAY-1996.
 PD
 XX
 XX 30-OCT-1995; 95WO-US14085.
 PF
 XX
 XX 04-NOV-1994; 94US-0334179.
 PR
 XX
 XX (PROC) PROCTER & GAMBLE CO.
 PA
 XX
 XX Nohno T, Rosenbaum JS;
 PI

XX WPI; 1996-251762/25.
 DR
 XX Isolated bone morphogenic protein receptor kinase protein - used to
 PT determine if a test cpd. is capable of binding to, or is
 PT (ant)agonist of BMP receptor kinase protein transcription
 XX
 XX Example 9; Page 30; 87pp; English.
 PS
 XX
 XX This primer corresponds to the sequence of plasmid pSPORT1/N89-5,
 CC and is the reverse complement of nucleotides 968-948 in a gene
 CC encoding mouse bone morphogenetic protein (BMP) receptor type-II
 CC kinase-3 (BRK-3), which induces cellular differentiation in response
 CC to BMP. The primer may be used with anchor primer AAT27231 in 5'-rapid
 CC amplification of cDNA ends (5'-RACE), to isolate 600 bp at the 5'-end
 CC of the BRK-3 gene, by PCR amplification of NIH3T3 mouse embryo
 CC fibroblast (ATCC CRL 1658) cDNA. To reduce background, a 2nd round
 CC of PCR is performed using nested primer AAT27232 with the anchor
 CC primer. This results in isolation of 600 bp from the 5'-end of the
 CC gene, which may be used to assemble a full-length clone (AAT27225) by
 CC PCR using primers AAT27223-4 and cloning in phage Bluescript-II-SK(-)
 CC in Escherichia coli DH5-alpha. The receptor and antibodies against
 CC it may be used in diagnostic assays for BMP disorders, or in
 CC therapy to bind or scavenge BMPs. In addition, expression of the
 CC BRK-3 gene along with a reporter gene under the control of a
 CC hormone-responsive element in a cell culture may be used to screen
 CC compounds for BRK-agonist or -antagonist activity, by monitoring
 CC reporter gene expression.
 XX
 SQ Sequence 21 BP; 6 A; 4 C; 6 G; 5 T; 0 other;

Query Match 86.7%; Score 13; DB 17; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggcttatcttca 13
 |||||
 Db 18 TGGCTTATCTTCA 6

RESULT 6
 AAX94668
 ID AAX94668 standard; DNA; 20 BP.
 AC
 XX
 XX AAX94668;
 AC
 XX 13-SEP-1999 (first entry)
 DT
 XX
 XX PCR primer used to amplify an ORF of Chlamydia pneumoniae.
 DE
 XX
 XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope; PCR primer; ss.
 OS
 XX Synthetic.
 OS Chlamydia pneumoniae.
 XX
 XX WO9927105-A2.
 PN
 XX
 XX 03-JUN-1999.
 PD
 XX
 XX 20-NOV-1998; 98WO-IB01890.
 PF
 XX
 XX 04-NOV-1998; 98US-0107078.
 PR
 XX 21-NOV-1997; 97FR-0014673.
 PR
 XX
 XX (GEST) GENSET.
 PA
 XX
 XX Griffais R;
 PI
 XX WPI; 1999-357842/30.
 DR
 XX

CC nucleus. RIZ gene mutations may be implicated in various cancers such as
 CC melanoma, neuroblastoma, leukaemia and breast cancer, and so the RIZ
 CC gene may be used in gene therapy for these disorders. Since RIZ protein
 CC is implicated in cell cycle arrest, inhibition of RIZ activity may be
 CC useful in neurodegenerative disorder therapy e.g. for Parkinson's,
 CC Huntington's or Alzheimer's disease, paralysis or motor neurone
 CC disorders, or cardiac disorders e.g. heart disease, where the ability to
 CC induce neural/ cardiac tissue proliferation would be useful. The present
 CC sequence was used in the construction of vectors which were used to clone
 CC mammalian RIZ coding sequences.

XX SQ Sequence 27 BP; 9 A; 7 C; 9 G; 2 T; 0 other;

Query Match 82.7%; Score 12.4; DB 21; Length 27;

Best Local Similarity 92.9%; Pred. No. 6.2e+02;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tggcttatcttcag 14

Db 27 TGGCTTTCTTCAG 14

RESULT 9

AAA57918

ID AAA57918 standard; DNA; 30 BP.

XX AC AAA57918;

XX DT 10-OCT-2000 (first entry)

XX DE Streptococcus pyogenes M19 gene PCR primer, SEQ ID NO:39.

XX KW Multivalent hybrid M protein: group A streptococcus; serotype;

XX KW immunogenic; sero-specific antibody; streptococcal infection;

XX KW cross reactivity; vaccine; acute rheumatic fever; ARF;

XX KW rheumatic heart disease; streptococcal pharyngitis; strep throat;

XX KW pneumonia; PCR primer; ss.

XX OS Streptococcus pyogenes.

XX PN US6063386-A.

XX PD 16-MAY-2000.

XX PF 15-SEP-1997; 97US-0937271.

XX PR 16-SEP-1992; 92US-0945954.

XX PA (UYTE-) UNIV TENNESSEE RES CORP.

XX PI Lederer JW, Dale JB;

XX DR WPI; 2000-364475/31.

XX PT New immunogenic recombinant hybrid M protein comprising amino-terminal

XX PT peptide fragments of streptococcal M protein useful as vaccine against

XX PT rheumatic fever and infections leading to rheumatic fever -

XX PS Disclosure; Column 13-14; 62pp; English.

XX CC The invention relates to multivalent immunogenic hybrid group A

XX CC streptococcal M proteins comprising N-terminal peptide fragments of M

XX CC proteins that elicit opsonic antibodies against multiple serotypes

XX CC of group A streptococci (e.g., Streptococcus pyogenes). The antibodies

XX CC generated using the hybrid proteins are against one or more M protein

XX CC serotypes, where at least one of the serotypes is M1, M3, M18 or M19.

XX CC The invention also encompasses a recombinant DNA molecule comprising a

XX CC nucleotide sequence that encodes a multivalent hybrid M protein; and a

XX CC method for immunising a mammal against streptococcal infections,

XX CC comprising administering an immunogenic multivalent hybrid M protein to

XX CC the mammal. The multivalent hybrid M proteins are useful for eliciting

XX CC opsonic or protective antibodies to the M proteins of different

CC serotypes of group A streptococci, and may therefore be used as vaccines
 CC to protect against and control infection by type A streptococci. Type A
 CC streptococci are not only responsible for streptococcal pharyngitis
 CC (strep throat), forms of pneumonia and a condition resembling toxic
 CC shock, but are also involved in the development of acute rheumatic fever
 CC (ARF) and rheumatic heart disease. In a patient with ARF, antibodies
 CC formed during a group A streptococcal infection are also cross-reactive
 CC with heart tissue, which indicates that the streptococci and host tissue
 CC contain similar antigenic motifs. The new multivalent vaccines are
 CC capable of raising sero-specific antibodies against various serotypes of
 CC group A streptococci which are not cross-reactive with human heart
 CC tissue. Sequences AAA57903-A57904 and AAA57908-A57919 represent PCR
 CC primers used in the generation of DNA (AAA57893-A57902) encoding
 CC multivalent hybrid Streptococcus pyogenes M proteins (AAB03113-B03117,
 CC AAB03119-B03121, AAB03123-B03124).

XX SQ Sequence 30 BP; 3 A; 6 C; 12 G; 9 T; 0 other;

Query Match 82.7%; Score 12.4; DB 21; Length 30;

Best Local Similarity 92.9%; Pred. No. 6.3e+02;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 tggcttatcttcag 14

Db 10 tggcttatcttcag 23

RESULT 10

AAF76114/C

ID AAF76114 standard; DNA; 19 BP.

XX AC AAF76114;

XX DT 22-MAY-2001 (first entry)

XX DE CCR5/CCR2b PCR primer, SEQ ID:18, used to genotype HIV susceptibility.

XX KW CC chemokine receptor; beta chemokine receptor; CCR; human;

XX KW CCR5; CCR2; polymorphism; genotyping; HIV-1 transmission;

XX KW infection susceptibility; AIDS; acquired immunodeficiency syndrome;

XX KW disease progression; chromosome 3p21-22; PCR primer; ss.

XX OS Homo sapiens.

XX PN WO200112857-A2.

XX PD 22-FEB-2001.

XX PF 11-AUG-2000; 2000WO-US22255.

XX PR 12-AUG-1999; 99US-0148530.

XX PA (UABR-) UAB RES FOUND.

XX PI Tang J, Kaslow RA;

XX DR WPI; 2001-211235/21.

XX PT Surveying CC beta chemokine receptor (CCR) genotypes in population,

XX PT involves amplifying genomic DNA of individuals with experimental and

XX PT control primer combinations, size-separating amplicons and determining

XX PS CCR genotype -

XX CC Claim 1; Page 42; 118pp; English.

XX CC The invention relates to a method of surveying the CC (beta) chemokine

XX CC receptor (CCR) genotypes in a population. The method is particularly

XX CC applied to the human CCR5 and CCR2 genes located on chromosome 3p21-22,

XX CC which encode co-receptors for HIV-1. The method involves obtaining

XX CC genomic DNA samples from a representative number of individuals within

XX CC a population; combining each sample with experimental and control primer

XX CC combinations to produce primer-annealed DNA; amplifying the DNA to

CC produce amplicons; separating the amplicons by size; determining the CCR
 CC genotype based upon the presence of CCR alleles; and compiling the
 CC genotypes determined. The method is particularly applied to the human
 CC CCR5 and CCR2 genes, which encode co-receptors for HIV-1. Polymorphisms
 CC in these genes are associated with a variation in the susceptibility of
 CC an individual to infection by HIV-1, or with a variation in the disease
 CC progression of AIDS after infection. The invention specifically claims
 CC the experimental PCR primers AAF76098-AAF76112, and the control PCR
 CC primers AAF76113-AAF76114 for surveying CCR5 and CCR2b genotypes.
 CC The method of the invention fulfills a longstanding need for the
 CC development of a rapid and informative genotyping strategy that can be
 CC readily applied to analyse CCR5, CCR2 and related genetic variants,
 CC and to evaluate the relationship of each genotype to HIV transmission
 CC and disease progression. The present sequence represents a human
 CC CCR5/CCR2b control PCR primer for use in the method of the
 CC invention.
 CC
 CC Sequence 19 BP; 6 A; 4 C; 5 G; 4 T; 0 other;

Query Match 80.0%; Score 12; DB 22; Length 19;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ggcttatcttca 13
 |||||
 DB 15 GGCTTATCTTCA 4

RESULT 11
 AAV84128/c
 ID AAV84128 standard; DNA; 21 BP.
 XX
 AC AAV84128;
 XX
 DT 15-MAR-1999 (first entry)
 XX
 DE HIV-1 co-receptor CCR5-delta32 variant PCR primer delta-32 inv.
 XX
 KW HIV-1; CCR5; CCR5-delta32; co-receptor; infection; diagnosis; AIDS;
 KW gene therapy; human; PCR; primer; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9854317-A1.
 XX
 PD 03-DEC-1998.
 XX
 PF 29-MAY-1998; 98WO-EP03437.
 XX
 PR 30-MAY-1997; 97US-0048057.
 XX
 PA (MOND-) FOND MONDIALE RECH & PREVENTION SIDA.
 XX
 PI Arenzana Siesdedos F, Beretta A, Braun J, Quillient C;
 XX
 DR WPI; 1999-059835/05.
 XX
 PT New CCR5 variant protein of the HIV-1 co-receptor - useful in
 PT developing resistance of CCR5-expressing cells to HIV-1 infection
 XX
 PS Claim 7; Page 8; 55pp; English.
 XX
 CC This is the nucleotide sequence of a claimed DNA primer, designated
 CC delta-32 inv, which can be used for the amplification of HIV-1
 CC CCR5 co-receptor nucleic acids. It corresponds to nucleotide
 CC 955 of wild-type CCR5 cDNA (see AAV84126). 4 pairs of primers (see
 CC AAV84127-34) are provided. These can be used in claimed methods for
 CC detecting and identifying the presence of a CCR5 variant in a
 CC sample. Primer pair delta-32 inv and delta-32 sens (see AAV84127)
 CC are particularly useful in detecting the presence of variant
 CC CCR5-delta32 (see AAV84159) that contains a 32 bp deletion when

CC compared to the wild-type allele. Individuals homozygous for the
 CC variant allele show resistance to HIV-1 infection. The detection
 CC of CCR5 variants may be used to identify individuals at lower risk
 CC of infection relative to the general population who, if infected,
 CC may exhibit slower progression to AIDS.
 XX
 SQ Sequence 21 BP; 7 A; 6 C; 5 G; 3 T; 0 other;

Query Match 80.0%; Score 12; DB 20; Length 21;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 ggcttatcttca 13
 |||||
 DB 13 GGCTTATCTTCA 2

RESULT 12
 AAV08563/c
 ID AAV08563 standard; DNA; 21 BP.
 XX
 AC AAV08563;
 XX
 DT 12-FEB-1999 (first entry)
 XX
 DE Primer for Transgene for epitope tagged TBP protein.
 XX
 KW TATA-box binding protein; epitope-tagged TBP; transcription complex; TAF;
 KW TBP associated factor; TAF-interaction factor; gene expression regulator;
 KW PCR primer; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN EP881288-A1.
 XX
 PD 02-DEC-1998.
 XX
 PF 26-MAY-1998; 98EP-0109516.
 XX
 PR 26-MAY-1997; 97EP-0108433.
 XX
 PA (FARH) HOECHST AG.
 XX
 PI Berglund E, Kirschbaum B, Meisterernst M, Polites G;
 XX
 DR WPI; 1999-001394/01.

Transgenic animal expressing epitope-tagged TATA-box binding protein
 - for isolating higher-order transcription complexes and specific
 factors that associate with the protein, useful as potential
 therapeutic agents
 XX
 PS Example 4; Page 11; 38pp; English.
 XX
 CC This sequence is a PCR primer for DNA encoding an epitope-tagged TATA-box
 CC binding protein (TBP) that is expressed by the transgenic non-human
 CC animals of the invention. The animals are used to produce TBP. TBP is
 CC used to isolate and characterise higher-order transcription complexes
 CC (from different tissue and cell types, optionally at different
 CC developmental stages). It is also used to identify new and/or specific
 CC TBP associated factors (TAFs, e.g. transcription factors, activators or
 CC inhibitors) and TAF-interaction factors, and to raise antibodies against
 CC TBP. The TAFs may be useful for regulating gene expression,
 CC e.g. disease-related genes, so are potential pharmaceuticals, also for
 CC identifying human analogues for use in drug screening. The antibodies are
 CC used for affinity purification of TBP and its complexes. TBP can isolate
 CC transcription complexes from a wide variety of different tissues and
 CC cells (contrast known methods that are limited to isolation from a
 CC particular cell type).
 XX
 SQ Sequence 21 BP; 6 A; 4 C; 8 G; 3 T; 0 other;

Query Match 78.7%; Score 11.8; DB 20; Length 21;
Best Local Similarity 86.7%; Pred. No. 1.3e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tggcttatcttcagt 15.
|||||
DB 19 TGGCTTAACTTCACT 5

RESULT 13
AA07298
ID AAX07298 standard; DNA: 22 BP.
AC AAX07298;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human SIV/HIV receptor Bob downstream PCR primer.
XX
KW SIV receptor; HIV receptor; Bob; translocation promoting agent;
KW Bonzo; G protein coupled receptor; Chemokine receptor; AIDS;
KW Infection; virus transmission; PCR; primer; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9903888-A1.
PD 28-JAN-1999.
XX
PF 17-JUL-1998; 98WO-US14857.
XX
PR 17-JUL-1997; 97US-0896155.
XX
PA (UYNV) UNIV NEW YORK STATE.
XX
XX Deng H, Kewalramani VN, Littman DR, Unutmaz D;
XX WPI; 1999-132164/11.
XX

XX New nucleic acid encoding the human translocation promoter Bonzo -
PT used to screen for potential agents for treating acquired immune
PT deficiency syndrome
XX
PS Example 1; Page 51; 97pp; English.
XX
CC This primer corresponds to a region 550 bp downstream of the ATG
CC initiation codon of human Bob (see AAX07292), a new HIV/SIV
CC traslocation promoter agent which acts with CD4 to serve as a
CC receptor for retroviral entry. It was used with an upstream
CC primer (see AAX07297) in the RT-PCR amplification of lymphocyte
CC RNA. Bob mRNA was detected in PHA-stimulated peripheral blood
CC mononuclear cells (PBMC), purified T cells, and weakly in
CC unstimulated PBMC. Expression was either minimal or absent in
CC macaques and B cells. Human, African green monkey and pig-tailed
CC macaque genes (see AAX07289-94) encoding 2 receptors, termed Bonzo
CC and Bob (see AAX97783-88), that are associated with retroviral entry
CC into cells have been identified. Diagnostic and therapeutic uses
CC of these novel translocation promoting agents are provided,
CC including a method for identifying potential agents for treatment
CC of SIV- and HIV-induced AIDS.
XX
SQ Sequence 22 BP; 6 A; 5 C; 4 G; 7 T; 0 other;

Query Match 78.7%; Score 11.8; DB 20; Length 22;
Best Local Similarity 86.7%; Pred. No. 1.3e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tggcttatcttcagt 15
|||||||
DB 19 TGGCTTAACTTCACT 5

Db 4 tggcttatcatcaat 18
RESULT 14
AAA66237/c
ID AAA66237 standard; DNA: 23 BP.
XX
AC AAA66237;
XX
DT 09-OCT-2000 (first entry)
XX
DE Dog genomic marker oligonucleotide sequence SEQ ID NO:99.
XX
KW Dog; genome; genomic marker; radiation hybrid map; identification;
KW chromosome location; gene marker; polymorphic microsatellite marker;
KW phenotype; behaviour; pedigree; ss.
XX
OS Canis familiaris.
XX
PN WO200029615-A2.
XX
PD 25-MAY-2000.
XX
PF 15-NOV-1999; 99WO-IB01907.
XX
PR 13-NOV-1998; 98US-0108193.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Galibert F, Andre C;
XX WPI; 2000-387821/33.
XX
PT New radiation hybrid map of the dog, Canine familiaris, genome, useful
PT for e.g. identifying genes implicated in phenotypic and behavioral
PT traits or in genetic diseases and for studying dog pedigrees -
XX
PS Claim 1; Page 57; 87pp; English.
XX
CC The present invention describes a radiation hybrid map of the dog
CC (Canine familiaris) genome comprising the genome location of a marker
CC selected from AAA66139 to AAA66942. The radiation hybrid map is useful
CC for identifying and localising dog genes, since it covers approximately
CC 80 % of the dog genome and provides a dense map integrating different
CC types (i.e. Type I and Type II) of markers. The map and the dog genome
CC markers (or complementary sequences) are especially useful to identify
CC genes responsible for phenotypic and behavioural traits in dogs, to
CC identify morbid genes, to analyse diseases and identify implicated genes
CC in such diseases and their alleles, and to study dog pedigrees. They
CC may also be useful for isolating corresponding human gene sequences
CC e.g. genes involved in genetic diseases.
XX
SQ Sequence 23 BP; 10 A; 6 C; 5 G; 2 T; 0 other;

Query Match 78.7%; Score 11.8; DB 21; Length 23;
Best Local Similarity 86.7%; Pred. No. 1.3e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tggcttatcttcagt 15
|||||
DB 17 TGGCTTGGTTTCACT 3

RESULT 15
AAQ52624
ID AAQ52624 standard; DNA: 30 BP.
XX
AC AAQ52624;
XX
DT 03-JUN-1994 (first entry)
XX
DE Primer deltaK3 huHGF.

XX Polymerase chain reaction; PCR; amplify; primer; PRK 5.1 huHGF;
KW human hepatocyte growth factor; huHGF; variant; E. coli; human;
KW fetal; kidney 293 cells; hepatocyte; ligand; receptor; activation; ss.
XX OS Synthetic.
XX PN WO9323550-A.
XX PD 25-NOV-1993.
XX PF 17-MAY-1993; 93WO-US04717.
XX PR 18-MAY-1992; 92US-0884811.
PR 18-MAY-1992; 92US-0885971.
PR 22-SEP-1992; 92US-0950572.
XX PA (GETH) GENENTECH INC.
XX PI Godowski PJ;
XX DR WPI; 1993-386581/48.
XX PT Activation of receptors, partic. of growth factors - using
PT conjugate of two ligands capable of binding to receptors, for
PT mimicking biological activity
XX Example 1; Page 49; 103pp; English.
XX CC The sequences given in AQ52616-28 are primers which were used in the
CC generation of PRK 5.1 human hepatocyte growth factor (huHGF) variant
CC molecules. The amplification products from these reactions were
CC cloned in E. coli cells and were then used to transfect human fetal
CC kidney 293 cells. The expressed proteins were used to study the
CC ability of huHGF to induce DNA synthesis of hepatocytes in primary
CC culture. The expressed proteins may be used as first and/or second
CC ligands in a conjugate for receptor activation. These conjugates
CC may be used to activate the respective receptors and thereby mimic
CC the biological activity of the corresponding native ligand.
XX SQ Sequence 30 BP; 7 A; 7 C; 4 G; 12 T; 0 other;

Query Match 78.7%; Score 11.8; DB 14; Length 30;
Best Local Similarity 86.7%; Pred. No. 1.3e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 tggcttatcttcagt 15
||| |||||
Db 7 tgacatatcttcagt 21

Search completed: October 2, 2001, 16:18:49
Job time: 15493 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:54 ; Search time 417.38 Seconds
(without alignments)
6.804 Million cell updates/sec

Title: US-09-757-100B-31

Perfect score: 15
Sequence: 1 taccagggtgagtct 15

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	15	100.0	15	3	US-09-377-310-31
2	15	100.0	20	3	US-09-377-310-11
c 3	12.4	82.7	21	1	US-08-136-118-7
c 4	12.4	82.7	22	2	US-08-117-952-668
5	11.8	78.7	20	4	US-09-101-8868-71
6	11.8	78.7	34	2	US-08-988-128-15
c 7	11.8	78.7	48	2	US-08-865-675-6
c 8	11.8	78.7	48	2	US-09-237-510-6
c 9	11.4	76.0	20	4	US-09-488-671-119
10	11.4	76.0	30	1	US-08-123-702-34
11	11	73.3	20	2	US-09-015-876-20
12	11	73.3	20	2	US-08-609-4438-53
c 13	11	73.3	21	4	US-09-273-565-86
14	11	73.3	28	1	US-08-456-103-4
15	11	73.3	28	2	US-09-067-908-4
16	11	73.3	45	1	US-09-015-876-21
17	10.8	72.0	20	3	US-09-257-799-49
18	10.8	72.0	20	3	US-08-920-919A-49
c 19	10.8	72.0	21	1	US-08-477-877B-78
20	10.8	72.0	21	1	US-08-477-877B-79
c 21	10.8	72.0	21	1	US-08-472-281A-78
22	10.8	72.0	21	1	US-08-472-281A-79
c 23	10.8	72.0	21	2	US-08-477-989B-78
24	10.8	72.0	21	2	US-08-477-989B-79
25	10.8	72.0	22	1	US-07-916-034-6
c 26	10.8	72.0	26	1	US-08-482-882-87
c 27	10.8	72.0	26	1	US-08-483-389-87

```

c 28 10.8 72.0 26 2 US-08-487-113D-87 Sequence 87, Appl
c 29 10.8 72.0 26 2 US-08-473-503-87 Sequence 87, Appl
c 30 10.8 72.0 26 2 US-08-483-932-87 Sequence 87, Appl
c 31 10.8 72.0 26 2 US-08-720-420A-87 Sequence 87, Appl
c 32 10.8 72.0 26 3 US-08-714-017-87 Sequence 87, Appl
c 33 10.8 72.0 26 3 US-08-475-680-87 Sequence 87, Appl
c 34 10.8 72.0 29 1 US-08-530-492-69 Sequence 69, Appl
c 35 10.8 72.0 29 4 US-08-906-517-69 Sequence 69, Appl
c 36 10.8 72.0 39 4 US-07-987-264-44 Sequence 44, Appl
c 37 10.8 72.0 40 1 US-08-482-882-100 Sequence 100, App
c 38 10.8 72.0 40 1 US-08-483-389-100 Sequence 100, App
c 39 10.8 72.0 40 2 US-08-487-113D-100 Sequence 100, App
c 40 10.8 72.0 40 2 US-08-473-503-100 Sequence 100, App
c 41 10.8 72.0 40 2 US-08-483-932-100 Sequence 100, App
c 42 10.8 72.0 40 2 US-08-720-420A-100 Sequence 100, App
c 43 10.8 72.0 40 3 US-08-714-017-100 Sequence 100, App
c 44 10.8 72.0 40 3 US-08-475-680-100 Sequence 100, App
c 45 10.8 72.0 42 1 US-08-482-882-68 Sequence 68, Appl

```

ALIGNMENTS

```

RESULT 1
US-09-377-310-31
: Sequence 31, Application US/09377310B
: Patent No. 6133031
: GENERAL INFORMATION:
: APPLICANT: Monia, Brett P.
: TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
: FILE REFERENCE: ISPH-0389
: CURRENT APPLICATION NUMBER: US/09/377,310B
: CURRENT FILING DATE: 1999-08-19
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 31
: LENGTH: 15
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: antisense sequence
US-09-377-310-31

```

```

Query Match 100.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. NO. 4.1;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 taccagggtgagtct 15
    |||
Db 1 taccagggtgagtct 15

```

```

RESULT 2
US-09-377-310-11
: Sequence 11, Application US/09377310B
: Patent No. 6133031
: GENERAL INFORMATION:
: APPLICANT: Monia, Brett P.
: TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
: FILE REFERENCE: ISPH-0389
: CURRENT APPLICATION NUMBER: US/09/377,310B
: CURRENT FILING DATE: 1999-08-19
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 11
: LENGTH: 20
: TYPE: DNA
: ORGANISM: Artificial Sequence

```

FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-11

Query Match 100.0%; Score 15; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 taccaggtagtct 15
| | | | | | | | | | | | | | | | | | | | | |
Db 3 taccaggtagtct 17

RESULT 3
US-08-136-118-7/c
; Sequence 7, Application US/08136118
; Patent No. 5580969
; GENERAL INFORMATION:
; APPLICANT: HOKE, Glenn D
; APPLICANT: BRADLEY, Matthews O
; APPLICANT: WILLIAMS, Taffy J
; APPLICANT: LEE, Che-Hung
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES DIRECTED
; AGAINST HUMAN ICAM-1
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Naval Medical Res. & Dev. Cmd.
; STREET: 8901 Wisconsin Ave.
; CITY: Bethesda
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136.118
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/918,259
; FILING DATE: 24-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Spevack, A. David
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: N.C. 75,776
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 295-6759
; TELEFAX: (202) 295-1022
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
US-08-136-118-7

Query Match 82.7%; Score 12.4; DB 1; Length 21;
Best Local Similarity 92.9%; Pred. No. 1.3e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 taccaggtagtc 14
| | | | | | | | | | | | | | | | | | | | | |
Db 18 TTCCAGGTGAGTC 5

RESULT 4

US-08-117-952-668/c
; Sequence 668, Application US/08117952
; Patent No. 5851760
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen A.
; APPLICANT: Smith, Michael W.
; TITLE OF INVENTION: METHOD FOR GENERATION OF SEQUENCE
; MAPS OF COMPLEX GENOMES
; NUMBER OF SEQUENCES: 797
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,952
; FILING DATE: 07-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/078,471
; FILING DATE: 15-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9423
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 668:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Oligonucleotide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-117-952-668

Query Match 82.7%; Score 12.4; DB 2; Length 22;
Best Local Similarity 92.9%; Pred. No. 1.3e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 acccaggtagtct 15
| | | | | | | | | | | | | | | | | | | | | |
Db 14 ATCCAGGTGAGTCT 1

RESULT 5
US-09-101-886B-71
; Sequence 71, Application US/09101886B
; Patent No. 6197507
; GENERAL INFORMATION:
; APPLICANT: BERG, THOMAS
; APPLICANT: TOLLERSRUD, OLE K
; APPLICANT: NILSEN, OIVIND
; TITLE OF INVENTION: GENETIC TEST FOR ALPHA-MANNOSIDOSIS
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BARBARA G. ERNST
; STREET: 555 13TH STREET, NW SUITE 701E
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/101.886B
FILING DATE: 29-JANUARY-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB97/00109
FILING DATE: 12-JAN-1997
ATTORNEY/AGENT INFORMATION:
NAME: ERNST, BARBARA G
REGISTRATION NUMBER: 30,377
REFERENCE/DOCKET NUMBER: 1181-240
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-783-6040
TELEFAX: 202-783-6031
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-09-101-886B-71

Query Match 78.7%; Score 11.8; DB 4; Length 20;
Best Local Similarity 86.7%; Pred. No. 2.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taccacagtgagtct 15
| | | | | | | | | | | | | | | | |
Db 4 TGCCACAGTGAGTGT 18

RESULT 6
US-08-988-128-15
Sequence 15, Application US/08988128
Patent No. 5994505
GENERAL INFORMATION:
APPLICANT: Ting, Jenny Pan-Yung
APPLICANT: Chin, Keh-Chin
TITLE OF INVENTION: No. 5994505el Forms of Class II MHC
TITLE OF INVENTION: Transactivator (CIITA)
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Myers Bigel Sibley & Sajovec, P.A.
STREET: P.O. Box 37428
CITY: Raleigh
STATE: No. 5994505th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/988,128
FILING DATE: 10-DEC-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Biswas, Sorojini J.
REGISTRATION NUMBER: 39,111
REFERENCE/DOCKET NUMBER: 5470-136
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 854-1400
TELEFAX: (919) 854-1401

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-988-128-15

Query Match 78.7%; Score 11.8; DB 2; Length 34;
Best Local Similarity 86.7%; Pred. No. 2.9e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taccacagtgagtct 15
| | | | | | | | | | | | | | | | |
Db 16 TACCACAGCTGTGTCT 30

RESULT 7
US-08-865-675-6/c
Sequence 6, Application US/08865675
Patent No. 5928869
GENERAL INFORMATION:
APPLICANT: Nadeau, James G.
APPLICANT: Pitner, James B.
APPLICANT: Linn, Carl P.
APPLICANT: Schram, James L.
TITLE OF INVENTION: DETECTION OF NUCLEIC ACIDS BY
TITLE OF INVENTION: FLUORESCENCE QUENCHING
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: R. J. Redrick, Becton Dickinson and Company
STREET: 1 Becton Drive
CITY: Franklin Lakes
STATE: NJ
COUNTRY: US
ZIP: 07417
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,675
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fugit, Donna R.
REGISTRATION NUMBER: 32,135
REFERENCE/DOCKET NUMBER: P-3746
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-865-675-6

Query Match 78.7%; Score 11.8; DB 2; Length 48;
Best Local Similarity 86.7%; Pred. No. 3e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taccacagtgagtct 15
| | | | | | | | | | | | | | | | |
Db 41 TACTCAGATGAGTCT 27

RESULT 8
US-09-237-510-6/c
Sequence 6, Application US/09237510
Patent No. 5958700

```

: GENERAL INFORMATION:
: APPLICANT: Nadeau, James G.
: APPLICANT: Pitner, James B.
: APPLICANT: Linn, Carl P.
: APPLICANT: Schram, James L.
: TITLE OF INVENTION: DETECTION OF NUCLEIC ACIDS BY
: TITLE OF INVENTION: FLUORESCENCE QUENCHING
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: R. J. Rodrick, Becton Dickinson and Company
: STREET: 1 Becton Drive
: CITY: Franklin Lakes
: STATE: NJ
: COUNTRY: US
: ZIP: 07417
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/237,510
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Fugit, Donna R.
: REGISTRATION NUMBER: 32,135
: REFERENCE/DOCKET NUMBER: P-3746
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 48 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-09-237-510-6

```

Query Match 78.7%; Score 11.8; DB 2; Length 48;
Best Local Similarity 86.7%; Pred. No. 3e+02;
Matches 13; Conservative 0; Mismatches 2; Indels

Qy	1	taccagg	15
Db	41	TACTCAGATGAGTCT	27

```

RESULT          9
US-09-488-671-119/c
; Sequence 119, Application US/09488671A
; Patent No. 6187545
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Madeline M. Butler
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Lex M. Cowser
; TITLE OF INVENTION: ANTISENSE MODULATION OF PEPCK-CYTOSOLIC EXPRESSION
; FILE REFERENCE: RTS-0123
; CURRENT APPLICATION NUMBER: US/09/488, 671A
; CURRENT FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 177
; SEQ ID NO 119
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-488-671-119

```

Query Match 76.0%; Score 11.4; DB 4; Length 20;
Best Local Similarity 92.3%; Pred. No. 4.8e+02;
Matches 12: Conservative 0; Mismatches 1; Indels

```

QY      3 cccaggtgagctct 15
      || |||||
Db      15 CCAAGGTGAGTCT 3

RESULT 10
US-08-123-702-34
; Sequence 34, Application US/08123702
; Patent No. 5604131
; GENERAL INFORMATION:
; APPLICANT: Wadsworth, Samuel
; APPLICANT: Snyder, Benjamin
; APPLICANT: Reddy, Vermuri, B.
; APPLICANT: Wel, Chamer
; TITLE OF INVENTION: A cDNA Genomic Hybrid Sequence Encoding APP770
; Patent No. 5604131
; TITLE OF INVENTION: Containing a Genomic DNA Insert of the KI and OX-2 Regions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,702
; FILING DATE: 17-SEPT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-123-702-34

```

```
Query Match      76.0%; Score 11.4; DB 1; Length 30;
Best Local Similarity 92.3%; Pred. No. 4.9e+02;
Matches 12; Conservative 0; Mismatches 1; Indels
```

Qy 3 ccaggtgagtct 15
|||||
Db 6 CCCAGGTAAGTCT 18

RESULT 11
US-09-015-876-20
; Sequence 20, Application US/09015876
; Patent No. H001825
; GENERAL INFORMATION:
; APPLICANT: ROMANO, JOSEPH W
; APPLICANT: SHUTLIEFF, ROXANNE
; APPLICANT: WILLIAMS, KIMBERLY G
; TITLE OF INVENTION: ISOTHERMAL TRANSCRIPTION BASED ASSAY
; TITLE OF INVENTION: FOR THE DETECTION OF HTLV I AND HTLV II
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: AKZO NOBEL PATENT DEPT.
; STREET: 1300 PICCARD DRIVE, SUITE 206
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/015.876
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KLESNER, SHARON N
; REGISTRATION NUMBER: 36,335
; REFERENCE/DOCKET NUMBER: T/98353
; TELEPHONE: 301-948-7400
; TELEFAX: 301-948-9751
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-015-876-20

```

```

Query Match 73.3%; Score 11; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 3 cccaggtgagt 13
   |||||
Db 10 CCCAGGTGAGT 20

```

RESULT 12

```

US-08-609-443B-53
; Sequence 53, Application US/08609443B
; Patent No. 5840693
; GENERAL INFORMATION:
; APPLICANT: ERIKSSON, Ulf
; APPLICANT: OLOFSSON, Birgitta
; APPLICANT: ALITALO, Kari
; APPLICANT: PAJUSOLA, Katri
; TITLE OF INVENTION: VASCULAR ENDOTHELIAL GROWTH FACTOR-B AND
; TITLE OF INVENTION: DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan, P.L.L.C.
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,443B
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/397,651
; FILING DATE: 01-MAR-1995
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/469,427
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/569,063
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 1064/41979C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-609-443B-53

```

```

Query Match 73.3%; Score 11; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 5 caggtgagtct 15
   |||||
Db 8 CAGGTGAGTCT 18

```

RESULT 13

```

US-09-273-565-86/C
; Sequence 86, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/273,565A
; CURRENT FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 09/055,699
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 08/820,170
; EARLIER FILING DATE: 1997-03-19
; EARLIER APPLICATION NUMBER: JP 63410/1996
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: JP 69163/1997
; EARLIER FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: C1 Primer
US-09-273-565-86

```

```

Query Match 73.3%; Score 11; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 5 caggtgagtct 15
   |||||
Db 13 CAGGTGAGTCT 3

```

RESULT 14

US-08-456-103-4
; Sequence 4, Application US/08456103
; Patent No. 5744304
; GENERAL INFORMATION:
; APPLICANT: Munford, Robert S.
; TITLE OF INVENTION: INFLAMMATION-INDUCED EXPRESSION OF A
; TITLE OF INVENTION: RECOMBINANT GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,103
; FILING DATE: Submitted Herewith
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UTSD:406/KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
US-08-456-103-4

Query Match 73.3%; Score 11; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 cccaggtgagt 13
Db 16 CCCAGGTGAGT 26

RESULT 15
US-09-067-908-4
; Sequence 4, Application US/09067908
; Patent No. 5851822
; GENERAL INFORMATION:
; APPLICANT: Munford, Robert S.
; TITLE OF INVENTION: INFLAMMATION-INDUCED EXPRESSION OF A
; TITLE OF INVENTION: RECOMBINANT GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/067,908
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/456,103
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: UTSD:406/KIT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
US-09-067-908-4

Query Match 73.3%; Score 11; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 cccaggtgagt 13
Db 16 CCCAGGTGAGT 26

Search completed: October 2, 2001, 16:03:55
Job time: 14599 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: October 2, 2001, 16:18:48 ; Search time 876.95 Seconds

(without alignments)
10.740 Million cell updates/sec

Title: US-09-757-100B-31

Perfect score: 15

Sequence: 1 taccagggtgagtct 15

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601:*

1: /SIDS8/gcgdata/geneseq/geneseq/NA1980.DAT:*

2: /SIDS8/gcgdata/geneseq/geneseq/NA1981.DAT:*

3: /SIDS8/gcgdata/geneseq/geneseq/NA1982.DAT:*

4: /SIDS8/gcgdata/geneseq/geneseq/NA1983.DAT:*

5: /SIDS8/gcgdata/geneseq/geneseq/NA1984.DAT:*

6: /SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT:*

7: /SIDS8/gcgdata/geneseq/geneseq/NA1986.DAT:*

8: /SIDS8/gcgdata/geneseq/geneseq/NA1987.DAT:*

9: /SIDS8/gcgdata/geneseq/geneseq/NA1988.DAT:*

10: /SIDS8/gcgdata/geneseq/geneseq/NA1989.DAT:*

11: /SIDS8/gcgdata/geneseq/geneseq/NA1990.DAT:*

12: /SIDS8/gcgdata/geneseq/geneseq/NA1991.DAT:*

13: /SIDS8/gcgdata/geneseq/geneseq/NA1992.DAT:*

14: /SIDS8/gcgdata/geneseq/geneseq/NA1993.DAT:*

15: /SIDS8/gcgdata/geneseq/geneseq/NA1994.DAT:*

16: /SIDS8/gcgdata/geneseq/geneseq/NA1995.DAT:*

17: /SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT:*

18: /SIDS8/gcgdata/geneseq/geneseq/NA1997.DAT:*

19: /SIDS8/gcgdata/geneseq/geneseq/NA1998.DAT:*

20: /SIDS8/gcgdata/geneseq/geneseq/NA1999.DAT:*

21: /SIDS8/gcgdata/geneseq/geneseq/NA2000.DAT:*

22: /SIDS8/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15	22	AAC65563 Human focal adhesi
2	15	100.0	20	22	AAC65543 Human focal adhesi
3	12.4	82.7	21	18	AAV58077 ICAM-1 antisense o
4	12.4	82.7	21	19	AAV38613 Human ICAM-1, E-se
5	12.4	82.7	22	16	AAQ82668 Chromosome 11 (loc
6	12.4	82.7	33	22	AAF98212 C neoforans strai
7	12	80.0	20	19	AAV28523 Blackcurrant rever
8	12	80.0	32	19	AAV99528 Human S1 receptor
9	11.8	78.7	34	19	AAV15506 Primer for adeno-a
10	11.8	78.7	48	20	AAV82665 Target oligonucleo
11	11.4	76.0	20	22	AAF62963 Mouse PEPCK-cytoso

C 12	11.4	76.0	25	21	AAA72481	Soybean chalcone r
C 13	11.4	76.0	46	11	AAO04939	Oligonucleotide ca
C 14	11	73.3	20	17	AAAT37941	VEGF-B186 exon 5 b
C 15	11	73.3	20	17	AAAT37943	VEGF-B167 exon 5 b
C 16	11	73.3	20	20	AAV63604	Human VEGF-B186 ex
C 17	11	73.3	20	22	AAV55942	Human eukaryotic i
C 18	11	73.3	21	18	AAV01904	Skeletal muscle sp
C 19	11	73.3	23	20	AAV08933	GATA-1 locus splic
C 20	11	73.3	28	19	AAV22641	Primer used to amp
C 21	10.8	72.0	20	19	AAV57078	Human Notch3 mutan
C 22	10.8	72.0	21	19	AAV62658	Humanised antibody
C 23	10.8	72.0	21	19	AAV62659	Humanised antibody
C 24	10.8	72.0	21	19	AAV22891	Humanised LO-CD2a
C 25	10.8	72.0	21	19	AAV22892	Humanised LO-CD2a
C 26	10.8	72.0	21	20	AAZ10191	PCR primer used to
C 27	10.8	72.0	21	20	AAZ10192	PCR primer used to
C 28	10.8	72.0	23	21	AAAL0299	HIV-1 LTR (long te
C 29	10.8	72.0	26	19	AAV56420	Human ICAM-R cDNA
C 30	10.8	72.0	26	20	AAV69192	Humanised ICR-1.1
C 31	10.8	72.0	26	21	AAA97162	PCR primer 8 used
C 32	10.8	72.0	26	21	AAA08308	Humanised ICR-1.1
C 33	10.8	72.0	26	21	AAZ24334	Humanised murine a
C 34	10.8	72.0	28	17	AAAT35702	3' VH primer WHIP3
C 35	10.8	72.0	29	19	AAV00357	Bacillus thuringie
C 36	10.8	72.0	29	22	AAAT73275	Oligonucleotide #6
C 37	10.8	72.0	30	18	AAAT44073	Human vascular end
C 38	10.8	72.0	30	20	AAV81019	Humanised 340 Vh a
C 39	10.8	72.0	30	21	AAAL5752	Human VEGF-3 3' PC
C 40	10.8	72.0	37	20	AAZ33426	Human HAP08 RACE P
C 41	10.8	72.0	37	21	AAZ40044	PCR primer FOGIVHF
C 42	10.8	72.0	40	19	AAV56432	Human ICAM-R cDNA
C 43	10.8	72.0	40	19	AAV54868	PCR primer 93 used
C 44	10.8	72.0	40	20	AAZ21887	Primer for antibod
C 45	10.8	72.0	40	20	AAV69200	Chimeric ICR-1.1 a

ALIGNMENTS

RESULT 1

AAC65563 AAC65563 standard; DNA; 15 BP.

XX AAC65563;

XX 12-FEB-2001 (first entry)

XX Human focal adhesion kinase antisense sequence #29.

DE Human; focal adhesion kinase; FAK; signal transduction; cancer;

KW embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphothioate; ss.

XX Homo sapiens.

XX US6133031-A.

XX 17-OCT-2000.

XX 19-AUG-1999; 99US-0377310.

XX 19-AUG-1999; 99US-0377310.

XX (ISIS-) ISIS PHARM INC.

XX Monia BP, Gaarde WA;

XX Chromosome 11 (loc

XX C neoforans strai

XX Blackcurrant rever

XX Human S1 receptor

XX Primer for adeno-a

XX Target oligonucleo

XX Mouse PEPCK-cytoso

PS Claim 15; Column 25; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX Sequence 15 BP; 3 A; 4 C; 4 G; 4 T; 0 other;

Query Match 100.0%; Score 15; DB 22; Length 15;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 taccacagtgagtct 15

Db 1 taccacagtgagtct 15

RESULT 2

AAC65543
 ID AAC65543 standard; DNA; 20 BP.

XX AAC65543;

DT 12-FEB-2001 (first entry)

XX Human focal adhesion kinase antisense sequence #9.

XX Human; focal adhesion kinase; FAK; signal transduction; cancer;

KW embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphorothioate; ss.

XX Homo sapiens.

XX US6133031-A.

XX 17-OCT-2000.

PF 19-AUG-1999; 99US-0377310.

PR 19-AUG-1999; 99US-0377310.

XX (ISIS-) ISIS PHARM INC.

PI Monia BP, Gaarde WA;

XX WPI; 2001-006141/01.

XX New antisense compounds for inhibiting focal adhesion kinase

PT expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

PS Claim 15; Column 23; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein
 CC is involved in integrin-mediated signal transduction, and is implicated
 CC in cancer, particularly colon, breast and oral tumours, embryonic
 CC development disorders, angiogenic disorders and wound healing. The
 CC antisense sequences, including the one shown here, can be used in the
 CC treatment of all of these.

XX Sequence 20 BP; 5 A; 4 C; 6 G; 5 T; 0 other;

Query Match

Best Local Similarity 100.0%; Score 15; DB 22; Length 20;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 taccacagtgagtct 15

Db 3 taccacagtgagtct 17

RESULT 3

AAT58077/c

ID AAT58077 standard; DNA; 21 BP.

XX AAT58077;

XX 18-MAR-1997 (first entry)

XX ICAM-1 antisense oligonucleotide #7.

XX Antisense; pre-mRNA; mature mRNA; vascular defect; tissue defect;

KW human intercellular adhesion molecule-1; ICAM-1; inflammation;

KW adult respiratory distress syndrome; multiple organ failure; GM1594;

XX septic shock; ss.

XX Synthetic.

XX US5580969-A.

XX 03-DEC-1996.

XX 24-JUL-1992; 92US-0918259.

XX 12-OCT-1993; 93US-0136118.

PR 24-JUL-1992; 92US-0918259.

XX (USNA) US SEC OF NAVY.

XX Bradley MO, Hoke GD, Lee C, Williams TJ;

XX WPI; 1997-033603/03.

XX Anti-sense oligo:nucleotide(s) for blocking ICAM-1 mRNA translation
 PT - for treating septic shock, adult respiratory distress syndrome
 PT etc.

XX Claim 1; Column 21; 16pp; English.

XX The sequences given in AAT58071-85 represent oligonucleotides which are
 CC antisense to sequences contained in the pre-mRNA or mature mRNA
 CC transcript of human intercellular adhesion molecule-1 (ICAM-1).
 CC These oligonucleotides may be used for treating septic shock and the
 CC manifestations of septic shock, e.g. inflammation, and vascular and
 CC tissue defects. They are also useful in the treatment of septic
 CC shock associated diseases, e.g. adult respiratory distress syndrome,
 CC multiple organ failure etc.

XX Sequence 21 BP; 5 A; 9 C; 5 G; 2 T; 0 other;

Query Match

Best Local Similarity 82.7%; Score 12.4; DB 18; Length 21;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 taccacagtgagtct 14

Db 18 TTCCACAGGTGATC 5

RESULT 4

AAV38613/c

ID AAV38613 standard; DNA; 21 BP.

XX AAV38613;

XX 13-OCT-1998 (first entry)

XX Human ICAM-1, E-selectin, VCAM-1 antisense oligonucleotide.

PT comprises two universal primers adapted for nucleic acid amplification
 protocol

PS Claim 6; Page 18; 88pp; English.

XX The present invention describes an assemblage comprising two primers,
 CC each of which can be used to amplify the intergenic spacer region IGS1
 CC from one of various strains of the yeast *Cryptococcus neoformans*. A
 CC number of primers and probes are provided, as are the sequences of the
 CC IGS1 for 91 *C. neoformans* strains. This is useful in the discrimination
 CC of pathogenic yeasts, and the sequences can be used to construct a
 CC database having the same purpose. The present sequence is a probe or
 CC primer described in the invention.

XX Sequence 33 BP; 7 A; 6 C; 13 G; 7 T; 0 other;

Query Match 82.7%; Score 12.4; DB 22; Length 33;
 Best Local Similarity 92.9%; Pred. No. 4.8e+02;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 acccaggtgagtct 15

Db 17 acccaggtgagtcat 30

RESULT 7

AAV28523

ID AAV28523 standard; DNA; 20 BP.

XX AAV28523;

DT 28-AUG-1998 (first entry)

XX Blackcurrant reversion virus RNA2 3' proximal fragment primer 5.

DE Blackcurrant reversion disease; BRV; rna2; diagnosis; Ribes; PCR;

XX primer; ss.

OS Synthetic.

OS Blackcurrant reversion virus.

XX WO9810100-A1.

XX 12-MAR-1998.

PF 01-SEP-1997; 97WO-FI00507.

XX 05-SEP-1996; 96FI-0003474.

XX (ABOA-) ABOATECH OY AB.

XX Latvala S, Lehto K, Lemmetty A, Susi P;

XX WPI; 1998-193642/17.

XX Diagnosing blackcurrant reversion disease in plants e.g.

PT blackcurrant - using reverse transcriptase-PCR with primers

PT amplifying cDNA fragment complementary to fragment of new

PT blackcurrant reversion virus

XX Example 2; Page 27; 38pp; English.

XX Primer 5 and primer 6 (see AAV28524) are designed to amplify a 481
 CC nucleotide DNA fragment initiating 265 nucleotides upstream of the
 CC polyA tail of RNA2 of a blackcurrant reversion virus (BRV) Finnish
 CC isolate genome (see AAV28525). They were used with primer 1,2 (see
 CC AAV28521-22), designed to amplify a 210 nucleotide fragment, to
 CC detect different isolates of BRV originating from widely different
 CC geographic locations. Primer pair 1,2 amplified the expected
 CC virus-specific fragments from all tested virus isolates, and the
 CC primer pair 5,6 amplified the expected virus-specific fragment from
 CC nearly all the tested virus isolates, indicating that the viral

CC sequences detecting by these primer pairs are well conserved in all
 CC isolates, including the common (E) and strong (R) forms of the
 CC reversion disease. The invention provides methods and kits
 CC for the diagnosis of blackcurrant reversion disease using RT-PCR.

XX Sequence 20 BP; 7 A; 5 C; 6 G; 2 T; 0 other;

Query Match 80.0%; Score 12; DB 19; Length 20;

Best Local Similarity 100.0%; Pred. No. 7.9e+02;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 acccaggtgagt 13

Db 8 acccaggtgagt 19

RESULT 8

AAAT99528

ID AAT99528 standard; DNA; 32 BP.

XX AC AAT99528;

DT 21-MAY-1998 (first entry)

XX Human ST receptor PCR primer.

DE ST receptor; heat stable toxin receptor; colorectal cancer; tumour;
 KW metastasis; diagnosis; human; PCR; primer; ss.

XX Synthetic.

OS Homo sapiens.

XX WO9742506-A1.

XX 13-NOV-1997.

PF 02-MAY-1997; 97WO-US07467.

XX 03-MAY-1996; 96US-0016564.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Carithers SL, Waldman SA;

XX WPI; 1998-008454/01.

XX Determining whether an individual has metastasised colorectal cancer
 PT cells and origin of tumour cells - by detecting presence of
 PT heat-stable toxin receptor on cells in a sample

XX Claim 14; Page 54; 62pp; English.

XX Claimed PCR primers (see AAT99462-T99531) hybridise to sequences that
 CC encode the extracellular domain of human heat-stable toxin (ST)
 CC receptor protein (see AAW37371), a highly specific marker for
 CC metastasised colorectal cancer cells. PCR using these primers
 CC provides specific and sensitive detection of human ST receptor
 CC expression. A specific primer pair comprises the primers given in
 CC AAT99528 and AAT99529. Claimed in vitro methods for determining whether
 CC or not (i) an individual has metastasised colorectal cancer cells,
 CC or (ii) a tumour cell is a colorectal cancer cell comprise the
 CC steps of examining a sample of extraintestinal tissue and/or body
 CC fluids or tumour cells from an individual to determine whether ST
 CC receptor protein is being expressed by cells in the sample.
 CC Expression is determined by immunoassay or by PCR using primers
 CC that selectively amplify ST receptor cDNA or mRNA (see also
 CC AAT97229).

XX Sequence 32 BP; 6 A; 11 C; 9 G; 6 T; 0 other;

Query Match

80.0%; Score 12; DB 19; Length 32;

```

Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 cccaggtgagtc 14
DB 12 cccaggtgagtc 23

RESULT 9
AAV15506
ID AAV15506 standard; DNA; 34 BP.
XX
AC AAV15506;
XX
DT 11-JUN-1998 (first entry)
XX
DE
XX
DE
XX
KW AAV; vector; rep 68/78; PCR primer; gene therapy; ss.
XX
OS Synthetic.
OS Adeno associated virus.
XX
PN WO9749824-A1.
XX
PD 31-DEC-1997.
XX
PF 24-JUN-1997; 97WO-DE01333.
XX
PR 24-JUN-1996; 96DE-1025188.
XX
PA (MEDI-) MEDIGENE GMBH.
PA (MEDI-) MEDIGENE AG.
XX
PI Bogedain C, Hallek M, Maass G;
XX
WPI; 1998-086641/08.
XX
System of adeno-associated viral vector and rep 68/78 sequence of
this virus - expression of which is delayed until replication of
viral DNA has started, provides large scale production of vectors
for gene therapy
XX
Example; Page 7; 20pp; German.
XX
The present sequence was used in the development of a novel system,
comprising an adeno-associated virus (AAV) vector containing a
foreign DNA, and the rep 68/78 sequence of AAV, the expression of
which is delayed. The components may be present in cis (in a single
agent) or in trans (in separate agents).
The system is used for production of AAV vectors, particularly for
gene therapy. The foreign DNA may encode a therapeutic protein,
e.g. interferon, interleukin, growth factor, coagulation factor or
metabolic enzyme, particularly one that increases the
immunogenicity of tumour cells, and/or a diagnostic protein. The
system produces AAV vectors on a large scale. It is based on the
discovery that the rep68 and 78 proteins interfere with replication
of AAV DNA, and that this interference is overcome by delaying
expression of these proteins.
XX
SQ Sequence 34 BP; 4 A; 13 C; 9 G; 8 T; 0 other;

Query Match 78.7%; Score 11.8; DB 19; Length 34;
Best Local Similarity 86.7%; Pred. No. 1.1e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taccaggtgagtc 15
DB 6 taccaggtgagtc 20

RESULT 10
AAV82665/c
ID AAV82665 standard; DNA; 48 BP.
XX
AC AAV82665;
XX
DT 25-FEB-1999 (first entry)
XX
DE
XX
DE
XX
Target oligonucleotide JN5.
KW
KW Detector oligonucleotide; hairpin structure; fluorescence;
KW primer extension; hybridization; signal primer; frameshift mutation;
KW ss.
XX
OS Synthetic.
XX
PN EP881302-A2.
XX
PD 02-DEC-1998.
XX
PF 28-MAY-1998; 98EP-0109682.
XX
PR 30-MAY-1997; 97US-0865675.
XX
PA (BECT ) BECTON DICKINSON & CO.
XX
PI Linn PC, Nadeau JG, Pitner BJ, Schram JL;
XX
WPI; 1999-001406/01.
XX
New detector oligo:nucleotide having base-paired region carrying
quenched dyes - where dyes become fluorescent when region is
unpaired, useful for detection and amplification of target nucleic
acid
XX
Example 2; Page 12; 20pp; English.
XX
Target oligonucleotides AAV82665-68 were designed to hybridise to the
detector oligonucleotide of the invention. The detector oligonucleotide
comprises a single-stranded target-binding region and an
intramolecularly base-paired secondary structure linked to two dyes
(donor and acceptor fluorophores). In the secondary structure,
fluorescence of the donor is quenched, but when it is linearised or
unfolded a change in some fluorescence parameter becomes detectable. A
target nucleic acid is detected by hybridizing it to a detector
oligonucleotide in which the secondary structure is 5' to the target
binding region, primer extension to produce a complementary strand using
the secondary structure as template, resulting in linearization or
unfolding of it and detecting a change in fluorescence. The detector
oligonucleotide are used to detect (by primer extension and
hybridization) and amplify (as signal primer) target sequences, e.g. for
detecting frameshift mutations.
XX
SQ Sequence 48 BP; 15 A; 8 C; 10 G; 15 T; 0 other;

Query Match 78.7%; Score 11.8; DB 20; Length 48;
Best Local Similarity 86.7%; Pred. No. 1.1e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 taccaggtgagtc 15
DB 41 TACTCAGATGAGTCT 27

RESULT 11
AAF62963/c
ID AAF62963 standard; DNA; 20 BP.
XX
AC AAF62963;
XX
DT 08-MAY-2001 (first entry)
XX
DE Mouse PEPCCK-cytosolic antisense oligonucleotide ISIS 113360.

```

XX Mouse; antiinflammatory; cytostatic; antisense gene therapy;
 KW phosphoenol pyruvate carboxykinase-cytosolic; PEPCK-cytosolic;
 XX infection; inflammation; tumour formation; phosphorothioate; ss.
 XX
 OS Mus musculus.
 XX
 PN US6187545-B1.
 XX
 PD 13-FEB-2001.
 XX
 PF 21-JAN-2000; 2000US-0488671.
 XX
 PR 21-JAN-2000; 2000US-0488671.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI McKay R, Butler MM, Wyatt J, Cowsert LM;
 XX
 DR WPI; 2001-190979/19.
 XX
 XX Antisense compound capable of modulating the expression of phosphoenol
 PT pyruvate carboxykinase-cytosolic, useful for preventing or delaying
 PT infection, inflammation or tumor formation -
 XX
 PS Example 17; Column 44; 64pp; English.
 XX
 CC The present sequence is one of a number of antisense compounds of up to
 CC 30 nucleobases in length that are capable of inhibiting the expression of
 CC phosphoenol pyruvate carboxykinase-cytosolic (PEPCK-cytosolic). The
 CC antisense compounds are useful for inhibiting the expression of
 CC PEPCK-cytosolic in cells or tissues. They are commonly used as research
 CC reagents and in diagnostics, e.g. to elucidate the function of particular
 CC genes. They are also useful for distinguishing between functions of
 CC various members of a biological pathway and for research use. The
 CC antisense compounds are also useful prophylactically, e.g. to prevent or
 CC delay infection, inflammation or tumour formation. The present sequence
 CC is a chimeric phosphorothioate oligonucleotide with 2'-MOE wings and a
 CC deoxy gap.
 XX
 SQ Sequence 20 BP; 5 A; 6 C; 5 G; 4 T; 0 other;

Query Match 76.0%; Score 11.4; DB 22; Length 20;
 Best Local Similarity 92.3%; Pred. No. 1.7e+03;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 ccacagtgagtct 15
 || |||||
 DB 15 CCAAGTGAGTCT 3

RESULT 12
 AAA72481/c
 ID AAA72481 standard; DNA; 25 BP.
 XX
 AC AAA72481;
 XX
 DT 19-DEC-2000 (first entry)
 XX
 DE Soybean chalcone reductase PCR primer, SEQ ID NO:65.
 XX
 KW Soybean; chalcone reductase; phenylpropanoid pathway;
 KW isoflavone synthase; isoflavonoid biosynthesis; defence response;
 KW attractant; repellent; signal compound; antibiotic; transgenic plant;
 KW transgenic seed; PCR primer; ss.
 XX
 OS Glycine max.
 XX
 PN WO200044909-A1.
 XX
 PD 03-AUG-2000.
 XX

PF 26-JAN-2000; 2000WO-US01772.
 XX
 PR 27-JAN-1999; 99US-0117769.
 PR 20-JUL-1999; 99US-0144783.
 PR 24-SEP-1999; 99US-0156094.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Fader GM, Jung W, McGonigle B, Odell JT, Yu X;
 XX
 PD WPI; 2000-543395/49.
 XX
 PT Nucleic acids encoding isoflavonoid synthases, useful for producing
 PT transgenic plants with increased production of isoflavonoids which are
 PT involved in defense against phytopathogenic microorganisms -
 XX
 PS Example 14; Page 52; 157pp; English.
 XX
 CC Sequences AAA72437, AAA72449-A72461, AAA72468 and AAA72474-A72477
 CC represent cDNAs encoding novel plant isoflavone synthases (AAB21052 and
 CC AAB21054- AAB21071). Also disclosed is the soybean cytochrome P450
 CC monooxygenase CYP93C1, encoded by a known sequence (AAA72444, NCBI No.
 CC 2739005), which was identified in the present invention as having
 CC isoflavone synthase activity. The invention also relates to expression
 CC constructs, transformed host cells, and transgenic plants and seeds
 CC comprising the novel cDNA sequences of the invention. The invention also
 CC encompasses methods of altering isoflavone synthase expression in a host
 CC cell, altering isoflavonoid levels in a plant, and identifying nucleic
 CC acids encoding other plant isoflavone synthases. Isoflavone synthase
 CC plays a key role in the biosynthesis of isoflavonoids. Isoflavonoids are
 CC a class of secondary metabolites mainly produced in leguminous plants by
 CC a branch of the phenylpropanoid pathway. Isoflavone synthase catalyses
 CC the first step in the branch of this pathway that commits metabolic
 CC intermediates to the synthesis of isoflavonoids. Isoflavonoids
 CC participate in the defence response of legumes against phytopathogenic
 CC microorganisms and are also involved in symbiotic relationships between
 CC the roots of legumes and rhizobial bacteria which eventually result in
 CC nodulation and nitrogen-fixation. They have also been shown to act as
 CC antibiotics, repellents, attractants, and signal compounds, and
 CC consumption of legume isoflavonoid is associated with health benefits
 CC in humans. The novel isoflavonoid synthases, and nucleic acids encoding
 CC them, are useful for altering the levels of isoflavonoids produced in
 CC legumes such as soybean, and for the production of isoflavonoids in
 CC plants which do not naturally produce them (e.g., maize, rice, wheat).
 CC Sequences AAA72478-A72481 represent PCR primers used in an
 CC exemplification of the invention to amplify DNA encoding soybean chalcone
 CC reductase (an enzyme that acts upstream of isoflavone synthase in the
 CC phenylpropanoid pathway).
 XX
 SQ Sequence 25 BP; 6 A; 7 C; 5 G; 7 T; 0 other;

Query Match 76.0%; Score 11.4; DB 21; Length 25;
 Best Local Similarity 92.3%; Pred. No. 1.7e+03;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 ccacagtgagtct 15
 |||||
 DB 17 CCCAGTGAGTTT 5

RESULT 13
 AAA04939/c
 ID AAA04939 standard; DNA; 46 BP.
 XX
 AC AAA04939;
 XX
 DT 24-OCT-1990 (first entry)
 XX
 DE Oligonucleotide carrying mutation for factor VIII gene.
 XX
 KW Human factor VIII analogue; ss.
 XX

PR 06-JUN-1995; 95US-0469427.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
PA (UYHE-) UNIV HELSINKI LICENSING LTD OY.
XX
XX Alitalo K, Eriksson U, Olofsson B, Pajusola K;
PI
XX WPI; 1996-412582/41.
DR
XX
XX Vascular endothelial growth factor VEGF-B proteins - useful to
PT accelerate angiogenesis in wound healing, also related nucleic acid
PT and antibodies for cancer diagnosis
XX
XX Example 7; Page 28; 107pp; English.
PS
XX AAT37933-T37946 represent the intron/exon boundaries for the human
CC vascular endothelial growth factor (VEGF) proteins of the invention (see
CC AAW04829, and AAW04831), which promote endothelial or mesodermal cell
CC proliferation. VEGF is also a glycosylated cationic dimer, and is
CC sometimes referred to as vascular permeability factor (VPF). VEGF has
CC diverse effects, depending on the specific biological context in which
CC it is found. VEGF is a potent endothelial cell mitogen, and directly
CC contributes to induction of angiogenesis in vivo by promoting
CC endothelial cell growth during normal embryonic development, wound
CC healing, and tissue regeneration/reorganisation. The VEGF proteins of
CC the invention share the angiogenic and other properties of VEGF, but are
CC distributed and expressed in tissues differently to VEGF. The proteins
CC can therefore be used to accelerate angiogenesis in wound healing.
CC Antibodies against the proteins can be used for inhibiting angiogenesis.
CC The antibodies can also be used diagnostically to quantitatively detect
CC VEGF-B. Primers complementary to the coding sequences for the proteins
CC of the invention can also be used to detect VEGF-B coding sequences.
CC Quantification of VEGF-B in cancer biopsy specimens may be useful as an
CC indicator of metastatic risk. VEGF-B expression in a cell can be
CC retarded using antisense sequences direct against the VEGF coding
CC sequences, this is especially useful in retarding VEGF expression in
CC tumour cells.
XX
SQ Sequence 20 BP; 5 A; 4 C; 6 G; 5 T; 0 other;

Query Match 73.3%; Score 11; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5 caggtgagctct 15
|||||
Db 8 caggtgagctct 18

Search completed: October 2, 2001, 16:18:48
Job time: 15492 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:03:54 ; Search time 417.38 Seconds

(without alignments)
6.804 Million cell updates/sec

Title: US-09-757-100B-30

Perfect score: 15

Sequence: 1 taagcagctgccatt 15

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	3	US-09-377-310-30
2	15	100.0	20	3	US-09-377-310-10
C 3	13.4	89.3	18	1	US-08-378-761A-45
C 4	13.4	89.3	18	1	US-08-485-286-45
C 5	12	80.0	29	3	US-08-826-964-4
6	11.8	78.7	50	1	US-08-171-389-486
7	11.8	78.7	50	1	US-08-171-389-487
8	11.8	78.7	50	1	US-08-123-936-486
9	11.8	78.7	50	1	US-08-123-936-487
10	11.8	78.7	50	2	US-08-475-228A-486
11	11.8	78.7	50	2	US-08-475-228A-487
12	11.8	78.7	50	3	US-08-482-080A-486
13	11.8	78.7	50	3	US-08-482-080A-487
14	11.8	78.7	50	5	PCT-US93-12388-486
15	11.8	78.7	50	5	PCT-US93-12388-487
16	11.4	76.0	20	2	US-08-832-658A-4
17	11.4	76.0	20	4	US-09-306-876A-5
18	11.4	76.0	20	4	US-09-306-876A-6
19	11.4	76.0	25	1	US-08-482-115B-26
20	11.4	76.0	25	2	US-08-472-802C-27
21	11.4	76.0	30	1	US-08-833-377-5
22	11.4	76.0	39	2	US-07-759-568-5
C 23	11.4	76.0	39	2	US-08-484-397A-24
C 24	11.4	76.0	39	2	US-08-834-655-15
C 25	11.4	76.0	39	3	US-09-363-574-15
C 26	11	73.3	22	2	US-08-479-614-21
C 27	11	73.3	28	1	US-08-479-852-12

28	11	73.3	28	1	US-08-479-852-64	Sequence 64, Appl
C 29	11	73.3	28	1	US-08-479-852-78	Sequence 78, Appl
30	11	73.3	28	1	US-08-479-852-92	Sequence 92, Appl
C 31	11	73.3	28	2	US-08-482-846-12	Sequence 12, Appl
32	11	73.3	28	2	US-08-482-846-64	Sequence 64, Appl
C 33	11	73.3	28	2	US-08-462-646-92	Sequence 78, Appl
34	11	73.3	28	2	US-08-462-646-92	Sequence 92, Appl
C 35	11	73.3	28	4	US-09-013-406-12	Sequence 12, Appl
36	11	73.3	28	4	US-09-013-406-64	Sequence 64, Appl
C 37	11	73.3	28	4	US-09-013-406-78	Sequence 78, Appl
38	11	73.3	28	4	US-09-013-406-92	Sequence 92, Appl
39	11	73.3	28	1	US-08-233-030-37	Sequence 37, Appl
40	11	73.3	35	1	US-08-343-682-12	Sequence 12, Appl
C 41	11	73.3	37	4	US-09-198-955A-21	Sequence 21, Appl
C 42	11	73.3	39	2	US-08-951-822-13	Sequence 13, Appl
C 43	11	73.3	39	4	US-09-173-043-28	Sequence 28, Appl
44	11	73.3	41	2	US-08-484-397A-19	Sequence 19, Appl
45	11	73.3	45	4	US-09-312-266-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-377-310-30
; Sequence 30, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-30

Query Match 100.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 9.6; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0;
Qy 1 taagcagctgccatt 15
Db 1 taagcagctgccatt 15

RESULT 2
US-09-377-310-10
; Sequence 10, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-10

Query Match 100.0%; Score 15; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
| | | | | | | | | | | | | | | | | |
Db 3 taagcagctgccatt 17

RESULT 3
US-08-378-761A-45/C
; Sequence 45, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/378,761A
; FILING DATE: 26-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-378-761A-45

Query Match 89.3%; Score 13.4; DB 1; Length 18;
Best Local Similarity 93.3%; Pred. No. 72;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
| | | | | | | | | | | | | | | | | |
Db 17 TAAGCAGCTGCAATT 3

RESULT 4
US-08-485-286-45/c
; Sequence 45, Application US/08485286
; Patent No. 5646026
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A

; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,286
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/378761
; FILING DATE: 26-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BORUCKI, ANDREA T
; REGISTRATION NUMBER: 33651
; REFERENCE/DOCKET NUMBER: 38272B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4846
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-485-286-45

Query Match 89.3%; Score 13.4; DB 1; Length 18;
Best Local Similarity 93.3%; Pred. No. 72;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
| | | | | | | | | | | | | | | | | |
Db 17 TAAGCAGCTGCAATT 3

RESULT 5
US-08-826-964-4/c
; Sequence 4, Application US/08826964
; Patent No. 6033663
; GENERAL INFORMATION:
; APPLICANT: Ketcham, Catherine M.
; TITLE OF INVENTION: Nucleic Acids Encoding GDP-Fucose
; TITLE OF INVENTION: Pyrophosphorylase
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,964
; FILING DATE: 09-APR-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/015,241
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 014137-0099100S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-826-964-4

Query Match 80.08; Score 12; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 agcagctgccat 14
|||||
Db 27 AGCAGCTGCCAT 16

RESULT 6
US-08-171-389-486
Sequence 486, Application US/08171389
Patent No. 5578444
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 486:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human T-cell Lymphotropic virus type
INDIVIDUAL ISOLATE: III (HIV-1)
US-08-171-389-486

Query Match 78.7%; Score 11.8; DB 1; Length 50;
Best Local Similarity 86.7%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
|||||
Db 24 TAAGCAGCTGCTTTT 38

RESULT 7
US-08-171-389-487
Sequence 487, Application US/08171389
Patent No. 5578444
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
APPLICANT: Fry, Kirk E.
TITLE OF INVENTION: Sequence-Directed DNA Binding
TITLE OF INVENTION: Molecules, Compositions and Methods
NUMBER OF SEQUENCES: 641
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,936
FILING DATE: 17-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081,070
FILING DATE: 22-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0175/G19P3

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 487:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Aids-associated retrovirus
INDIVIDUAL ISOLATE: (arv-2;proviral)
US-08-171-389-487

Query Match 78.7%; Score 11.8; DB 1; Length 50;
Best Local Similarity 86.7%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
|||||
Db 24 TAAGCAGCTGCTTTT 38

RESULT 8
US-08-123-936-486
Sequence 486, Application US/08123936
Patent No. 5726014
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
TITLE OF INVENTION: Screening Assay for the Detection of
DNA-Binding Molecules
NUMBER OF SEQUENCES: 640
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,936
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 486:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: RNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Human T-cell Lymphotropic virus type
INDIVIDUAL ISOLATE: III (HIV-1)
US-08-123-936-486

Query Match 78.7%; Score 11.8; DB 1; Length 50;
Best Local Similarity 86.7%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 taagcagctgccatt 15
|||||
Db 24 TAAGCAGCTGCTTTT 38

RESULT 9
US-08-123-936-487
Sequence 487, Application US/08123936
Patent No. 5726014
GENERAL INFORMATION:
APPLICANT: Edwards, Cynthia A.
APPLICANT: Cantor, Charles R.
APPLICANT: Andrews, Beth M.
APPLICANT: Turin, Lisa M.
TITLE OF INVENTION: Screening Assay for the Detection of
DNA-Binding Molecules
NUMBER OF SEQUENCES: 640
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genelabs Technologies, Inc.
STREET: 505 Penobscot Drive
CITY: Redwood City
STATE: CA
COUNTRY: USA
ZIP: 94063
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/123,936
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/996,783
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/723,618
FILING DATE: 27-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 4600-0075.32/G19P2
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 487:
SEQUENCE CHARACTERISTICS:
LENGTH: 50 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: Aids-associated retrovirus
INDIVIDUAL ISOLATE: (arv-2;proviral)
US-08-123-936-487

Query Match 78.7%; Score 11.8; DB 1; Length 50;

Best Local Similarity 86.7%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
|||||
Db 24 TAAGCAGCTGCTTTT 38

RESULT 10
US-08-475-228A-486
; Sequence 486, Application US/08475228A
; Patent No. 5869241
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 664
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,228A
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,070
; FILING DATE: 22-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stratford, Carol A.
; REGISTRATION NUMBER: 34,444
; REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 486:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human T-cell Lymphotropic virus type
; INDIVIDUAL ISOLATE: III (HIV-1)
US-08-475-228A-486

Query Match 78.7%; Score 11.8; DB 2; Length 50;
Best Local Similarity 86.7%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
|||||
Db 24 TAAGCAGCTGCTTTT 38

RESULT 11
US-08-475-228A-487
; Sequence 487, Application US/08475228A
; Patent No. 5869241
; GENERAL INFORMATION:
; APPLICANT: Edwards, Cynthia A.
; APPLICANT: Cantor, Charles R.
; APPLICANT: Andrews, Beth M.
; APPLICANT: Turin, Lisa M.
; APPLICANT: Fry, Kirk E.
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 664
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,228A
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/723,618
; FILING DATE: 27-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081,070
; FILING DATE: 22-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stratford, Carol A.
; REGISTRATION NUMBER: 34,444
; REFERENCE/DOCKET NUMBER: 4600-0175.21/G19P3D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 487:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Aids-associated retrovirus
; INDIVIDUAL ISOLATE: (arv-2;proviral)
US-08-475-228A-487

Query Match 78.7%; Score 11.8; DB 2; Length 50;
Best Local Similarity 86.7%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
|||||
Db 24 TAAGCAGCTGCTTTT 38

Qy 1 taagcagctgcatt 15
|||||
Db 24 TAAGCAGCTGCTTTT 38

```
RESULT 14
PCT-US93-12388-486
; Sequence 486, Application PC/TUS9312388
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12388
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 486:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Human T-cell Lymphotropic virus type
; INDIVIDUAL ISOLATE: III (HIV-1)
PCT-US93-12388-486

Query Match 78.7%; Score 11.8; DB 5; Length 50;
Best Local Similarity 86.7%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
Db 24 TAAGCAGCTGCTTTT 38

RESULT 15
PCT-US93-12388-487
; Sequence 487, Application PC/TUS9312388
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Sequence-Directed DNA Binding
; TITLE OF INVENTION: Molecules, Compositions and Methods
; NUMBER OF SEQUENCES: 641
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genelabs Technologies, Inc.
; STREET: 505 Penobscot Drive
```

```
; CITY: Redwood City
; STATE: CA
; COUNTRY: USA
; ZIP: 94063
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12388
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,936
; FILING DATE: 17-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/996,783
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0175.41/G19PCT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 487:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: Aids-associated retrovirus
; INDIVIDUAL ISOLATE: (arv-2;proviral)
PCT-US93-12388-487

Query Match 78.7%; Score 11.8; DB 5; Length 50;
Best Local Similarity 86.7%; Pred. No. 5.8e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
Db 24 TAAGCAGCTGCTTTT 38

Search completed: October 2, 2001, 16:03:54
Job time: 14598 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2001, 16:18:47 ; Search time 876.95 seconds
(without alignments)
10.740 Million cell updates/sec

Title: US-09-757-100B-30

Perfect score: 15

Sequence: 1 taagcagctgccatt 15

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 854978

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	15	100.0	15	22	AAC65562 Human focal adhesi
2	15	100.0	20	22	AAC65542 Human focal adhesi
3	14	93.3	22	20	AA336327 Sense primer used
4	12.4	82.7	27	20	AA201016 PCR primer for PGL
5	12.4	82.7	49	21	AAA64261 PCR primer for hum
6	12	80.0	29	18	AAT96994 Human GDP-fucose p
7	12	80.0	29	18	AAT97584 Human GDP-fucose p
8	11.8	78.7	21	22	AAT95258 Human gene single
9	11.8	78.7	27	18	AA67939 Human fittl VEGF re
10	11.8	78.7	30	19	AAV65645 HIV-1 promoter fra
11	11.8	78.7	46	14	AAQ49401 HIV-1 TATA region.

12	11.8	78.7	50	15	AAQ69736 Human T-cell lymph
13	11.8	78.7	50	15	AAQ69737 AIDS-associated re
14	11.8	78.7	50	18	AAT64198 Human T-cell lymph
15	11.8	78.7	50	18	AAT64199 Aids-associated re
16	11.8	78.7	50	19	AAV65660 HIV-1 TATA region
17	11.8	78.7	50	20	AAV17486 Test sequence from
18	11.8	78.7	50	20	AAV17487 Polymorphic fragme
19	11.6	77.3	29	21	AAA04041 IGF-I oligonucleot
20	11.4	76.0	15	22	AAFS3844 IGF-I oligonucleot
21	11.4	76.0	15	22	AAFS3845 IGF-I oligonucleot
22	11.4	76.0	15	22	AAFS3846 Human PRO1555 reve
23	11.4	76.0	18	21	AAC58087 Primer #141 used i
24	11.4	76.0	18	22	AAFS4534 PCR primer used to
25	11.4	76.0	20	20	AAV63648 PCR primer used to
26	11.4	76.0	20	20	AAV63648 Human epidermal gr
27	11.4	76.0	20	20	AAV63648 EGFR oligonucleoti
28	11.4	76.0	20	22	AAV7717 EGFR oligonucleoti
29	11.4	76.0	20	22	AAV7718 Native Pseudomonas
30	11.4	76.0	22	16	AAQ98175 WMV2 Mar6 oligonuc
31	11.4	76.0	22	20	AAV7812 Human c-myc cDNA P
32	11.4	76.0	24	21	AAV74420 Soybean chlorophyl
33	11.4	76.0	25	17	AAT35689 RNA component of h
34	11.4	76.0	25	17	AAT10302 Human telomerase R
35	11.4	76.0	25	20	AAZ07280 Antisense oligonuc
36	11.4	76.0	30	19	AAV63648 Antisense oligonuc
37	11.4	76.0	30	20	AAZ23630 Human clone 28-1 t
38	11.4	76.0	30	20	AAV72932 Human KM-102-deriv
39	11.4	76.0	31	14	AAV33374 KHCV cDNA fragment
40	11.4	76.0	31	20	AAQ06130 Human biallelic po
41	11.4	76.0	39	14	AAQ43806 Human IL-8 recepto
42	11.4	76.0	39	19	AAV63629 PCR primer pRDS1 u
43	11.4	76.0	39	21	AAA09435 M. alpina delta-6
44	11.4	76.0	41	11	AAQ06571 Sequence at juncti
45	11.4	76.0	41	19	AAV51159 Maize polymorphic

ALIGNMENTS

RESULT 1

AAC65562 AAC65562 standard; DNA; 15 BP.

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

XX AC AAC65562;

PS Claim 15; Column 25; 30pp; English.

XX The present invention describes a number of phosphorothioate antisense

CC sequences to the human focal adhesion kinase (FAK) protein. This protein

CC is involved in integrin-mediated signal transduction, and is implicated

CC in cancer, particularly colon, breast and oral tumours, and is implicated

CC development disorders, angiogenic disorders and wound healing. The

CC antisense sequences, including the one shown here, can be used in the

CC treatment of all of these.

XX

SQ Sequence 15 BP; 4 A; 4 C; 3 G; 4 T; 0 other;

Query Match 100.0%; Score 15; DB 22; Length 15;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 15

Db ||||||||||||

Db 1 taagcagctgccatt 15

RESULT 2

AAC65542

ID AAC65542 standard; DNA; 20 BP.

AC AAC65542;

XX

DT 12-FEB-2001 (first entry)

XX

DE Human focal adhesion kinase antisense sequence #8.

XX

DE Human; focal adhesion kinase; FAK; signal transduction; cancer;

KW embryonic development disorder; angiogenic disorder; wound healing;

KW antisense; phosphorothioate; ss.

XX

OS Homo sapiens.

XX

PN US6133031-A.

PN

PD 17-OCT-2000.

PD

XX

PF 19-AUG-1999; 99US-0377310.

XX

PR 19-AUG-1999; 99US-0377310.

XX

PA (ISIS-) ISIS PHARM INC.

XX

PI Monia BP, Gaarde WA;

XX

DR WPI; 2001-006141/01.

XX

PT New antisense compounds for inhibiting focal adhesion kinase

PT expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -

XX

PS Example 2; Column 23; 30pp; English.

XX

CC The present invention describes a number of phosphorothioate antisense

CC sequences to the human focal adhesion kinase (FAK) protein. This protein

CC is involved in integrin-mediated signal transduction, and is implicated

CC in cancer, particularly colon, breast and oral tumours, embryonic

CC development disorders, angiogenic disorders and wound healing. The

CC antisense sequences, including the one shown here, can be used in the

CC treatment of all of these.

XX

SQ Sequence 20 BP; 5 A; 4 C; 5 G; 6 T; 0 other;

Query Match 100.0%; Score 15; DB 22; Length 20;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 15

Db ||||||||||||

Db 1 taagcagctgccatt 15

RESULT 3

AAX36327/c

ID AAX36327 standard; DNA; 22 BP.

XX

AC AAX36327;

XX

DT 21-JUL-1999 (first entry)

XX

DE Sense primer used to amplify human FAK cDNA.

XX

KW Human; tumorigenicity; glycosyltransferase;

KW malignancy; brain cancer cell; protein glycosylation; glioma;

KW meningioma; brain tumour; FAK; PCR primer; ss.

XX

OS Synthetic.

XX

PN WO9924584-A1.

PN

PD 20-MAY-1999.

PD

XX

PF 12-NOV-1998; 98WO-US24224.

XX

PR 12-NOV-1997; 97US-0969437.

XX

PA (NEUR-) NEUROTHERAPEUTICS.

XX

PI Moskal JR, Yamamoto H;

XX

DR WPI; 1999-327411/27.

XX

PT Altering tumorigenicity and malignancy of brain cancer cells

XX

PS Example 3; Page 34; 83pp; English.

XX

CC PCR primers AAX36327-28 were used to amplify human FAK cDNA, in the

CC course of the invention. The specification describes a method for

CC altering the tumorigenicity or malignancy of brain cancer cells by

CC changing the activity of glycosyltransferase in the cell so that

CC glycosylation of cellular proteins is modified. The method is applied

CC to glioma or meningioma, for prevention or treatment of brain tumours.

CC Measuring the level of glycosyltransferase expression in brain cells

CC is used to detect or predict their tumorigenicity.

XX

SQ Sequence 22 BP; 4 A; 7 C; 5 G; 6 T; 0 other;

Query Match 93.3%; Score 14; DB 20; Length 22;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 taagcagctgccatt 14

Db ||||||||||||

Db 14 TAAGCAGCTGCCAT 1

RESULT 4

AAZ01016/c

ID AAZ01016 standard; DNA; 27 BP.

XX

AC AAZ01016;

XX

DT 27-SEP-1999 (first entry)

XX

DE PCR primer for PGI gene exon border.

XX

KW PGI gene; biallelic marker; PCR primer; PGI-related biallelic marker;

KW cancer; prostate cancer; diagnosis; therapy; prostate specific antigen;

KW PSA; human; ss.

```

XX Synthetic.
OS Homo sapiens.
XX WO9932644-A2.
XX 01-JUL-1999.
XX 22-DEC-1998; 98WO-IB02133.
XX 09-SEP-1998; 98US-0099658.
XX 22-DEC-1997; 97US-0996306.
XX (GEST ) GENSET.
XX Blumenfeld M, Bouqueleret L, Chumakov I, Cohen D;
PI WPI; 1999-405178/34.
XX Use of a prostate cancer associated gene and biallelic markers
PT derived from it
XX Example 8; Page 261; 385pp; English.
XX The invention relates to a mammalian PGI gene and protein, and a set of
CC PGI biallelic markers. The PGI polynucleotide and biallelic markers are
CC used in a hybridisation assay, a sequencing assay, or in an
CC allele-specific amplification assay for determining the identity of a
CC nucleotide at a PGI-related biallelic marker. The methods can be used to
CC detect and to assess the risk of developing cancer or prostate cancer.
CC Early-stage diagnosis of prostate cancer relies on prostate specific
CC antigen (PSA) dosage. However, the effectiveness of this is limited due
CC to its inability to discriminate between malignant and non-malignant
CC affections of the organ. A need exists for both a reliable diagnostic
CC procedure which would enable early-stage diagnosis, and for preventative
CC and curative treatments of the disease. The PGI gene can be used for
CC detection of prostate cancer, and the risk of developing it in the
CC future, and can also be used to determine therapies for the disease.
XX Sequence 27 BP; 6 A; 4 C; 6 G; 11 T; 0 other;
SQ
Query Match 82.7%; Score 12.4; DB 20; Length 27;
Best Local Similarity 92.9%; Pred. No. 7.7e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 taagcagctgccat 14
Db 24 TAAGCATCTGCCAT 11
RESULT 5
AAA64261/c
ID AAA64261 standard; DNA; 49 BP.
XX AAA64261;
XX 20-DEC-2000 (first entry)
XX PCR primer for human fibronectin collagen-binding domain cDNA.
XX Fibronectin; collagen-binding domain; sustained release; gene therapy;
KW physiologically active polypeptide; topical retention; PCR primer;
KW tissue regeneration; ss.
XX Homo sapiens.
XX WO200049159-A1.
XX 24-AUG-2000.
XX 21-FEB-2000; 2000WO-JP00964.
XX

```

```

PR 19-FEB-1999; 99JP-0041913.
PR 01-NOV-1999; 99JP-0311364.
XX (TERU ) TERUMO CORP.
XX Ishikawa T, Kitajima T;
XX WPI; 2000-565375/52.
XX Collagen-binding active polypeptide for use in an agent for enabling
PT topical retention or sustained release of a physiologically active
PT peptide or physiological activity-imparting agent comprises a
PT fibronectin peptide.
XX Example 1; Page 124; 135pp; English.
XX PCR primers AAA64261-62 were used to amplify cDNA encoding a human
CC fibronectin collagen-binding domain. The amplified sequence was
CC used to construct a collagen-binding physiologically active polypeptide.
CC This polypeptide comprises a peptide from fibronectin ligated to a
CC physiologically active peptide. The polypeptides are used in an agent
CC for enabling topical retention or sustained release of a physiologically
CC active peptide or physiological activity-imparting agent. They may
CC be used in gene therapy and in tissue regeneration.
XX Sequence 49 BP; 12 A; 16 C; 11 G; 10 T; 0 other;
SQ
Query Match 82.7%; Score 12.4; DB 21; Length 49;
Best Local Similarity 92.9%; Pred. No. 8.2e+02;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 taagcagctgccat 14
Db 31 TAAACAGCTGCCAT 18
RESULT 6
AAT96994/c
ID AAT96994 standard; DNA; 29 BP.
XX AAT96994;
XX 22-APR-1998 (first entry)
XX Human GDP-fucose pyrophosphorylase (GDPFPP) amplifying 5' primer.
XX Guanosine 5-diphospho- beta L-fucose pyrophosphorylase; human; GDPFPP;
KW carbohydrate; donor substrate; GDP-fucose; PCR primer; ss.
XX Synthetic.
OS Homo sapiens.
XX WO9737682-A1.
XX 16-OCT-1997.
XX 09-APR-1997; 97WO-US05876.
XX 10-APR-1996; 96US-0015241.
XX (CYTE-) CYTEL CORP.
PA (KETC/) KETCHAM C M.
XX WPI; 1997-512414/47.
XX Nucleic acids encoding GDP-fucose pyrophosphorylase - useful to
PT synthesise specific carbohydrate structures e.g. to investigate role
PT of carbohydrate(s) in cell surface recognition
XX Example 2; Page 24; 33pp; English.
XX This primer is used for the PCR amplification of guanosine 5-diphospho-
CC

```

CC beta L-fucose (GDP-fucose) pyrophosphorylase (GDPFPP) cDNA. Cells can be
 CC genetically engineered to contain the GDPFPP nucleic acids and produce
 CC the protein. GDPFPP enzymes can be used to synthesise carbohydrate
 CC molecules of defined structures, useful in investigating the role of
 CC carbohydrates as recognition elements on cell surfaces. They are
 CC especially useful in producing donor substrates (e.g. GDP-fucose) in
 CC reactions of a glycosyl transferase (e.g. fucosyltransferase) with the
 CC substrate, an acceptor sugar and a divalent metal cation to allow
 CC formation of glycosidic linkages adding a saccharide to a substrate
 CC saccharide. The proteins can also be used to produce antibodies or
 CC antisera useful to characterise, detect and isolate proteins
 CC cross-reacting with the GDPFPP protein.
 CC Note: The specification refers to claimed GDPFPP nucleotide and
 CC amino acid sequences, neither of which are given in the specification.
 XX
 SQ Sequence 29 BP; 6 A; 6 C; 10 G; 7 T; 0 other;

Query Match 80.0%; Score 12; DB 18; Length 29;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 agcagctgccat 14
 Db 27 AGCAGCTGCCAT 16

RESULT 7
 AAT97584/c
 ID AAT97584 standard; DNA; 29 BP.

XX AC AAT97584;

XX DT 14-APR-1998 (first entry)

XX DE Human GDP-fucose pyrophosphorylase cDNA 5' PCR primer.

XX KW GDP-fucose pyrophosphorylase; GDPFPP; human; oligosaccharide;
 KW carbohydrate; PCR; primer; ss.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN W09737683-A1.

XX PD 16-OCT-1997.

XX PF 10-APR-1997; 97WO-US05968.

XX PR 09-APR-1997; 97US-0831590.

XX PR 10-APR-1996; 96US-0015241.

XX PA (CYTE-) CYTEL CORP.

XX PI Ketcham CM;

XX DR WPI; 1997-512415/47.

XX PT GDP-fucose pyrophosphorylase and related coding sequences - useful
 PT to synthesise GDP-fucose as substrate for producing specific
 PT carbohydrate structures, e.g. to study cell surface recognition

XX PS Example 2; Page 24; 40pp; English.

XX CC A 5' primer (AAT97584) and a 3' primer (AAT97585) were used in the PCR
 CC amplification of human GDP-fucose pyrophosphorylase (GDPFPP) cDNA
 CC (see AAT97583) from Epstein-Barr virus-transformed B lymphoblastoid
 CC cell line JY, and were designed to incorporate EcoRV sites at the
 CC 5' and 3' termini of the amplified sequence. The amplified GDPFPP
 CC cDNA was subsequently incorporated into expression vectors.
 CC Recombinant human GDPFPP was expressed in COS, NSO and Spodoptera
 CC frugiperda Sf9 cells. The claimed enzyme (see AAT97581) is useful in
 CC the synthesis of carbohydrate structures.

XX SQ Sequence 29 BP; 6 A; 6 C; 10 G; 7 T; 0 other;

Query Match 80.0%; Score 12; DB 18; Length 29;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 agcagctgccat 14

Db 27 AGCAGCTGCCAT 16

RESULT 8
 AAF95258/c
 ID AAF95258 standard; DNA; 21 BP.

XX AC AAF95258;

XX DT 06-JUN-2001 (first entry)

XX DE Human gene single nucleotide polymorphism #19.

XX KW Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
 KW polymorphism; vascular disease; coronary artery disease; forensics;
 KW myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
 KW pulmonary embolism; paternity test; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT Variation replace(11,T)
 FT /*tag= a
 FT /standard_name= "single nucleotide polymorphism"

XX PN W0200118250-A2.

XX PD 15-MAR-2001.

XX PF 07-SEP-2000; 2000WO-US24503.

XX PR 10-SEP-1999; 99US-0153357.

XX PR 26-JUL-2000; 2000US-0220947.

XX PR 16-AUG-2000; 2000US-0225724.

XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Lander ES, Gargill M, Ireland JS, Bolk S, Daley GQ, McCarthy JJ;

XX DR WPI; 2001-226749/23.

XX PT Nucleic acids comprising single nucleotide polymorphisms, useful in
 PT applications such as forensics, paternity testing, medicine, genetic
 PT analysis and phenotype correlations to diseases such as diabetes and
 PT atherosclerosis -

XX PS Examples; Page 48; 242pp; English.

XX CC The present invention provides a method of diagnosing a vascular disease
 CC in an individual, involving determining the sequence at various
 CC polymorphic sites within the human thrombospondin 1 and thrombospondin 4
 CC genes. The sequences at a number of polymorphic sites are also provided
 CC in the specification. In particular, the method can be used in the
 CC diagnosis of atherosclerosis, myocardial infarction, coronary heart
 CC disease, stroke, peripheral vascular diseases, venous thromboembolism
 CC and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
 CC useful in forensics, paternity testing, genetic analysis and phenotype
 CC correlations to diseases. The present sequence is an example of one of
 CC the human gene SNPs shown in the specification.

XX SQ Sequence 21 BP; 6 A; 7 C; 4 G; 4 T; 0 other;

Query Match 78.7%; Score 11.8; DB 22; Length 21;
 Best Local Similarity 86.7%; Pred. No. 1.6e+03;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 taagcagctgccatt 15
 || |||||
 Db 17 TAGGCAGCTGCATT 3

RESULT 9
 AAX67939/c
 ID AAX67939 standard; RNA; 27 BP.
 XX
 AC AAX67939;
 XX
 DT 28-JUL-1999 (first entry)
 XX
 DE Human flt1 VEGF receptor hammerhead ribozyme #665.
 XX
 KW Vascular endothelial growth factor receptor; VEGF receptor; flt-1;
 KW flk-1; KDR; hammerhead ribozyme; hairpin ribozyme; cleavage;
 KW tumour angiogenesis; psoriasis; rheumatoid arthritis; ocular disease;
 KW fms-like tyrosine kinase 1; kinase insert domain containing receptor;
 KW foetal liver kinase 1; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9715662-A2.
 XX
 PD 01-MAY-1997.
 XX
 XX 25-OCT-1996; 96WO-US17480.
 XX
 PR 11-JAN-1996; 96US-0584040.
 PR 26-OCT-1995; 95US-0005974.
 XX
 PA (CHIR) CHIRON CORP.
 PA (RIBO-) RIBOZYME PHARM INC.
 XX
 PI Escobedo J, McSwiggen J, Pavco P, Stinchcomb D;
 XX
 XX WPI; 1997-259017/23.
 XX
 PT Nucleic acid molecule modulating VEGF receptor(s) gene expression or
 PT mRNA stability - useful for treating e.g. tumour angiogenesis,
 PT psoriasis, rheumatoid arthritis, etc., in a human patient
 XX
 PS ClaIm 9; Page 66; 218pp; English.

The present invention describes nucleic acid molecules which modulate the synthesis, expression and/or stability of a mRNA encoding 1 or more receptors of vascular endothelial growth factor (VEGF). A patient (preferably human) having a condition associated with the level of the fms-like tyrosine kinase 1 (flt-1), kinase insert domain containing receptor (KDR) and/or foetal liver kinase 1 (flk-1) (e.g. tumour angiogenesis, ocular diseases, psoriasis and rheumatoid arthritis) can be treated by administering the nucleic acid molecule or the expression vector to the patient. AAX67275 to AAX75752 represent specific examples of nucleic acid molecules from the present invention.

Sequence 27 BP; 9 A; 4 C; 9 G; 4 U; 1 other;

Query Match 78.7%; Score 11.8; DB 18; Length 27;
 Best Local Similarity 86.7%; Pred. No. 1.6e+03;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 taagcagctgccatt 15
 | | |||||
 Db 15 TCATCAGCTGCCATT 1

RESULT 10

AAV65645
 ID AAV65645 standard; DNA; 30 BP.

XX
 AC AAV65645;

XX
 DT 16-DEC-1998 (first entry)

XX
 DE HIV-1 promoter fragment used as a target sequence.

XX
 KW Polyamide; DNA-binding ligand; modulation; gene expression; CBP; IM;
 KW minor groove transcription factor protein; MGFPP; TFIIA; adenocarcinoma;
 KW transcription inhibitor; pathogen; binding affinity; zinc finger protein;
 KW carboxamide binding pair; N-methylimidazole carboxamide; bacteria; fungi;
 KW 5S ribosomal RNA gene; internal control region; virus; protozoa; HIV-1;
 KW human immune deficiency virus; cancer; DNA capture; genomic sequencing;
 KW her-2/neu oncogene; gene transcription; DNA cleavage; ds.

XX
 OS Human immunodeficiency virus type 1.

XX
 FH Key Location/Qualifiers
 FT misc_binding 6..11
 FT FT /*tag= a
 FT FT /note= "six basepair sequence bound by the
 FT FT polyamide"

FT FT TATA_signal 13..17
 FT FT /*tag= b
 FT FT /note= "TFIID binding site"

FT FT misc_binding 18..23
 FT FT /*tag= c
 FT FT /note= "six basepair sequence bound by the
 FT FT polyamide"

XX
 PN WO9835702-A1.

XX
 PD 20-AUG-1998.

XX
 XX 11-FEB-1998; 98WO-US02444.

XX
 XX 21-JUL-1997; 97WO-US12722.

PR 14-FEB-1997; 97US-0038384.

PR 14-FEB-1997; 97US-0038394.

PR 21-APR-1997; 97US-0853022.

XX
 XX (CALY) CALIFORNIA INST OF TECHNOLOGY.

PA (SCRI) SCRIPPS RES INST.

XX
 PI Baird EE, Dervan PB, Gottesfeld JW, Mosier DE;

XX
 DR WPI; 1998-506287/43.

XX
 PT Modulating expression of genes with polyamide(s) specific for
 PT region near the binding site for transcription factor - for
 PT inhibiting replication of pathogen, especially human immune
 PT deficiency virus and for treating cancers associated with expression
 PT of the her-2/neu oncogene

XX
 PS Example 3; Fig 17; 113pp; English.

XX
 CC This represents a promoter fragment of the HIV-1 genome that is
 CC used as a target sequence. The invention provides polyamide DNA-binding
 CC ligands for modulating expression of cellular and viral genes. One method
 CC comprises identifying a unique target DNA sequence adjacent to the
 CC binding site of a minor groove transcription factor protein (MGTFP),
 CC choosing a polyamide having subnanomolar affinity for the target DNA and
 CC treating the target DNA with the polyamide as transcription inhibitor.
 CC Methods (1) for inhibiting replication of a pathogen by administering a
 CC transcription-inhibiting polyamide; (2) improving binding affinity of
 CC polyamide, selected for an identified viral DNA target by replacing a
 CC carboxamide binding pair (CBP) that does not include N-methylimidazole
 CC carboxamide (IM) with a CBP comprising paired beta-alanine (beta)
 CC residues; (3) inhibiting binding of the zinc finger protein TFIIA to

CC the 5S ribosomal RNA gene internal control region; (4) for treating
 CC adenocarcinoma of ovary, endometrium, breast, fallopian tubes and cervix
 CC with the polyamide are also provided. Method (1) is applied to viruses,
 CC bacteria, fungi and protozoa, especially human immune deficiency virus
 CC (HIV)-1, both therapeutically and for treating blood cells in vitro.
 CC Method (4) is especially used against cancers that overexpress the
 CC her-2/neu oncogene. The polyamide can also be used for diagnosis of
 CC disease, very generally as therapeutic agents for any disease involving
 CC cellular or viral gene transcription, for genomic sequencing, for DNA
 CC capture and for DNA cleavage (oxidative or by light). The polyamide has
 CC excellent specificity and very high affinity for the target DNA, which
 CC are specific for particular genes, and are cell permeable.
 XX
 SQ Sequence 30 BP; 7 A; 6 C; 7 G; 10 T; 0 other;

Query Match 78.7%; Score 11.8; DB 19; Length 30;
 Best Local Similarity 86.7%; Pred. No. 1.7e+03;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
 ||| ||| ||| ||| |||
 Db 15 taagcagctgctttt 29

RESULT 11
 AAQ49401
 ID AAQ49401 standard; cDNA; 46 BP.

XX AC AAQ49401;

XX DT 27-APR-1994 (first entry)

XX DE HIV-1 TATA region.

XX KW TATA modulating factor; TMF; transcription; TATA box; promoter; HIV-1;
 KW human immunodeficiency virus-1; short arm; human chromosome 3; p12-p21;
 KW translocation; cancer; ss.

OS Human immunodeficiency virus-1.

PN WO9320106-A.

PD 14-OCT-1993.

XX PF 31-MAR-1993; 93WO-US03077.

XX PR 02-APR-1992; 92US-0862025.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PI Gaynor RB, Wu F;

XX DR WPI; 1993-336836/42.

XX PT New protein cellular factor - capable of binding double stranded
 PT HIV-1 tata region and activating gene expression of HIV-1TR

XX PS Claim 3; Page 48; 75pp; English.

XX CC This sequence represents the TATA region of the HIV-1 LTR from -46
 CC to -1. This region is bound by TATA modulating factor (TMF). TMF
 CC is a protein of mol. wt. 123-130 kD which activates transcription in
 CC most genes, esp. in human immunodeficiency virus-1 (HIV-1) by binding
 CC to the TATA box region of the promoter. TMF is encoded by the short
 CC arm of human chromosome 3 in the region p12-p21 which is often
 CC involved in translocations in patients having lung and other types
 CC of cancer.

XX SQ Sequence 46 BP; 8 A; 12 C; 11 G; 15 T; 0 other;

Query Match 78.7%; Score 11.8; DB 14; Length 46;

Best Local Similarity 86.7%; Pred. No. 1.7e+03;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15
 ||| ||| ||| ||| |||
 Db 22 taagcagctgctttt 36

RESULT 12

AAQ69736

ID AAQ69736 standard; DNA; 50 BP.

XX AC AAQ69736;

XX DT 06-MAR-1995 (first entry)

XX DE Human T-cell lymphotropic virus type III (HIV-1), target region.

XX KW DNA protein-binding assay; test sequence; screening sequence;

XX KW promoter; target; TATA box; Herpes Simplex Virus; HSV;

XX KW origin of replication; UL9; transcription factor; TFIID; ds.

XX OS Synthetic.

XX PN WO9414980-A.

XX PD 07-JUL-1994.

XX PF 20-DEC-1993; 93WO-US12388.

XX PR 23-DEC-1992; 92US-0996783.

XX PR 17-SEP-1993; 93US-0123936.

XX PA (GENE-) GENELABS TECHNOLOGIES INC.

XX PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;

XX DR WPI; 1994-234711/28.

XX PT Sequence-directed DNA-binding molecules - useful in
 PT pharmaceuticals and as molecular reagents

XX PS Claim 28; Page 455; 587pp; English.

XX CC A DNA protein-binding assay is provided, useful for screening
 CC libraries of synthetic or biological cpds. for their ability
 CC to bind DNA test sequences. The assay is versatile in that any
 CC number of test sequences can be tested by placing the test sequence
 CC adjacent to a defined protein-binding screening sequence. Binding
 CC of mols. to these test sequences changes the binding characteristics
 CC of the protein mol. to its cognate binding sequence. When such a mol.
 CC binds the test sequence, the equilibrium of the DNA:protein complexes
 CC is disturbed, generating changes in the concentration of free DNA probe.
 CC One application of this method is to eucaryotic general transcription
 CC factors (e.g. TFIID), where the target region is typically selected
 CC from DNA sequences adjacent to the binding site for the eucaryotic
 CC transcription factor. Numerous exemplary test sequences are given:
 CC the sequences in AAQ69251-731 and AAQ69850 correspond to promoter
 CC targets (typically, TATA box-contg. sites) for human genes and the
 CC sequences in AAQ69732-849 correspond to promoter targets for viral genes.
 CC The test sequences may also be randomly generated. DNA:protein
 CC interaction may be used for screening purposes, e.g. the Herpes Simplex
 CC Virus (HSV) origin of replication and UL9 (see AAQ69851-52, AAQ69865 and
 CC AAQ69891).

XX SQ Sequence 50 BP; 9 A; 14 C; 13 G; 14 T; 0 other;

Query Match 78.7%; Score 11.8; DB 15; Length 50;
 Best Local Similarity 86.7%; Pred. No. 1.8e+03;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgccatt 15

Db 24 taagcagctgctttt 38
|||||

RESULT 13
AAQ69737
ID AAQ69737 standard; DNA; 50 BP.

XX AAQ69737;
XX 06-MAR-1995 (first entry)
XX
XX AIDS-associated retrovirus (arv-2; proviral), target region.
XX
XX DNA protein-binding assay; test sequence; screening sequence;
XX promoter; target; TATA box; Herpes Simplex Virus; HSV;
XX origin of replication; UL9; transcription factor; TFIID: ds.
XX
XX Synthetic.

XX WO9414980-A.
XX 07-JUL-1994.
XX
XX 20-DEC-1993; 93WO-US12388.
XX
XX 23-DEC-1992; 92US-0996783.
XX 17-SEP-1993; 93US-0123936.
XX (GENE-) GENELABS TECHNOLOGIES INC.

XX Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX WPI; 1994-234711/28.
XX
XX Sequence-directed DNA-binding molecules - useful in
XX pharmaceuticals and as molecular reagents
XX
XX Claim 28; Page 455; 587pp; English.

XX A DNA protein-binding assay is provided, useful for screening
XX libraries of synthetic or biological cpds. for their ability
XX to bind DNA test sequences. The assay is versatile in that any
XX number of test sequences can be tested by placing the test sequence
XX adjacent to a defined protein-binding screening sequence. Binding
XX of mols. to these test sequences changes the binding characteristics
XX of the protein mol. to its cognate binding sequence. When such a mol.
XX binds the test sequence, the equilibrium of the DNA:protein complexes
XX is disturbed, generating changes in the concentration of free DNA probe.
XX One application of this method is to eucaryotic general transcription
XX factors (e.g. TFIID), where the target region is typically selected
XX from DNA sequences adjacent to the binding site for the eucaryotic
XX transcription factor. Numerous exemplary test sequences are given:
XX the sequences in AAQ69251-731 and AAQ69850 correspond to promoter
XX targets (typically, TATA box-contg. sites) for human genes and the
XX sequences in AAQ69732-849 correspond to promoter targets for viral genes.
XX The test sequences may also be randomly generated. DNA:protein
XX interaction may be used for screening purposes, e.g. the Herpes Simplex
XX Virus (HSV) origin of replication and UL9 (see AAQ69851-52, AAQ69865 and
XX AAQ69891).

XX Sequence 50 BP; 8 A; 13 C; 13 G; 16 T; 0 other;

Query Match 78.7%; Score 11.8; DB 15; Length 50;
Best Local Similarity 86.7%; Pred. No. 1.8e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgctttt 15
|||||

Db 24 taagcagctgctttt 38

RESULT 14
AAT64198
ID AAT64198 standard; DNA; 50 BP.

XX AAT64198;
XX
XX 17-MAR-1997 (first entry)
XX
XX Human T-cell lymphotropic virus type-III (HIV-1) TFIID binding site.
XX
XX Duplex DNA; target region; binding characteristic; DNA binding protein;
XX TFIID; transcription factor; binding site; inhibition; enhance;
XX cancer; inherited genetic disorder; ds.
XX
XX Human lymphotropic virus type III.
XX OS
XX US5578444-A.
XX
XX 26-NOV-1996.

XX 27-JUN-1991; 91US-0723618.
XX
XX 20-DEC-1993; 93US-0171389.
XX 27-JUN-1991; 91US-0723618.
XX 23-DEC-1992; 92US-0996783.
XX 17-SEP-1993; 93US-0123936.
XX (GENE-) GENELABS TECHNOLOGIES INC.

XX Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX WPI; 1997-020402/02.
XX
XX Altering binding characteristics of DNA binding proteins to duplex
XX DNA - by attaching specific small cpd. to target region close to the
XX protein's binding site, useful in treatment of viral disease, cancer
XX etc
XX
XX Claim 6; Column 347-348; 264pp; English.

XX The sequences given in AAT63713-4312 represent duplex DNA's which act
XX as target regions in the method of the invention. The method for
XX altering the binding characteristics of a DNA-binding protein to duplex
XX DNA comprises contacting the duplex DNA with a small molecule which
XX binds sequence-specifically to a target region, where, when the small
XX molecule is bound to the target region, it is adjacent to, but not
XX overlapping by more than 4 bp, a binding site for a DNA-binding protein.
XX The small molecule is added at a concentration effective to alter the
XX binding of the DNA binding protein, pref. TFIID, to its binding site on
XX the duplex DNA. The binding of the small molecule may inhibit or
XX enhance the binding of the DNA-binding protein to its binding site. The
XX compounds isolated using this method are potentially useful as
XX therapeutic agents for treatment of any disease which involves a
XX specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
XX The method is suitable for screening large biological or chemical
XX libraries and allows determination of sequence-specific and relative
XX affinities of known DNA-binding agents for different DNA sequences.
XX The design of these duplex DNA's allows a single DNA:protein interaction
XX to be used for screening sequence-specific, or preferential, DNA binding
XX proteins that recognise almost any possible sequence (see also AAT49539-
XX 74).

XX Sequence 50 BP; 9 A; 14 C; 13 G; 14 T; 0 other;

Query Match 78.7%; Score 11.8; DB 18; Length 50;
Best Local Similarity 86.7%; Pred. No. 1.8e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 taagcagctgctttt 15
|||||

Db 24 taagcagctgctttt 38

RESULT 15
AAT64199
ID AAT64199 standard; DNA; 50 BP.
XX
AC AAT64199;
XX
DT 17-MAR-1997 (first entry)
XX
DE Aids-associated retrovirus proviral arv2 TFIIID binding site.
XX
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIIID; transcription factor; binding site; inhibition; enhance;
KW cancer; inherited genetic disorder; ds.
XX
OS Aids-associated retrovirus.
XX
PN US5578444-A.
XX
PD 26-NOV-1996.
XX
PF 27-JUN-1991; 91US-0723618.
XX
PR 20-DEC-1993; 93US-0171389.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX WPI; 1997-020402/02.
XX
DR Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc
XX
PS Claim 6; Column 347-348; 264pp; English.
XX
CC The sequences given in AAT63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or
CC enhance the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also AAT49539-
CC 74).
XX
SQ Sequence 50 BP; 8 A; 13 C; 13 G; 16 T; 0 other;

Search completed: October 2, 2001, 16:18:48
JOB time: 15492 sec

Query Match 78.7%; Score 11.8; DB 18; Length 50;
Best Local Similarity 86.7%; Pred. NO. 1.8e+03;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 taagcagctgcatt 15
|||||
Db 24 taagcagctgctttt 38

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic --nucleic search, using sw model

Run on: October 2, 2001, 16:03:53 ; Search time 417.38 Seconds
(without alignments)
6.804 Million cell updates/sec

Title: US-09-757-100B-23

Perfect score: 15

Sequence: 1 gcggggtcacagtgg 15

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 460742

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	100.0	15	3	US-09-377-310-23
2	15	100.0	20	3	US-09-377-310-3
3	13.4	89.3	31	3	US-09-282-996-18
4	11.8	78.7	31	2	US-08-732-612-6
5	11.8	78.7	39	2	US-08-732-612-11
6	11.4	76.0	18	2	US-09-161-015-17
7	11.4	76.0	22	1	US-08-445-289B-25
8	11	73.3	25	2	US-08-859-998-856
9	11	73.3	31	1	US-07-959-946-8
10	11	73.3	31	1	US-08-333-577-8
11	11	73.3	31	5	PCT-US92-08634-8
12	10.8	72.0	18	2	US-09-166-203-9
13	10.8	72.0	24	1	US-08-206-176-10
14	10.8	72.0	24	1	US-08-371-001-2
15	10.8	72.0	24	2	US-08-756-506-9
16	10.8	72.0	24	3	US-09-131-956-9
17	10.8	72.0	24	5	PCT-US96-00331-2
18	10.8	72.0	29	1	US-08-271-880A-50
19	10.8	72.0	29	2	US-08-910-408-50
20	10.8	72.0	29	3	US-09-249-215-50
21	10.8	72.0	40	3	US-09-135-639-8
22	10.8	72.0	40	3	US-09-135-639-10
23	10.8	72.0	41	6	5190871-2
24	10.8	72.0	48	3	US-08-488-551B-802
25	10.8	72.0	50	4	US-09-315-886C-5
26	10.4	69.3	17	4	US-09-275-680-11
27	10.4	69.3	20	4	US-09-428-219-86

28 10.4 69.3 21 1 US-08-171-718-51 Sequence 51, Appl
29 10.4 69.3 21 1 US-08-468-580-50 Sequence 50, Appl
30 10.4 69.3 21 3 US-08-478-087-51 Sequence 51, Appl
31 10.4 69.3 21 3 US-08-905-359A-4 Sequence 4, Appl
32 10.4 69.3 21 4 US-09-353-556-4 Sequence 4, Appl
33 10.4 69.3 21 4 US-08-643-212-72 Sequence 72, Appl
34 10.4 69.3 21 5 PCT-US95-03731-50 Sequence 50, Appl
35 10.4 69.3 23 4 US-09-011-745-18 Sequence 18, Appl
36 10.4 69.3 26 2 US-08-859-998-969 Sequence 969, Appl
37 10.4 69.3 31 4 US-08-828-533-36 Sequence 36, Appl
38 10.4 69.3 39 6 5520913-18 Patent No. 5520913
39 10.4 69.3 43 4 US-08-250-802-34 Sequence 34, Appl
40 10.4 69.3 43 4 US-08-250-802-35 Sequence 35, Appl
41 10.4 69.3 43 5 PCT-US92-07916-34 Sequence 34, Appl
42 10.4 69.3 43 5 PCT-US92-07916-35 Sequence 35, Appl
43 10.2 68.0 17 1 US-08-379-078-471 Sequence 471, Appl
44 10.2 68.0 17 5 PCT-US93-00977-84 Sequence 84, Appl
45 10.2 68.0 18 3 US-09-344-579-40 Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-09-377-310-23
; Sequence 23, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO*23
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-23

Query Match 100.0%; Score 15; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gcggggtcacagtgg 15
|||||
Db 1 gcggggtcacagtgg 15

RESULT 2
US-09-377-310-3
; Sequence 3, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence

; OTHER INFORMATION: /note= "downstream primer
; OTHER INFORMATION: designated murf4"
US-08-732-612-11

Query Match 78.7%; Score 11.8; DB 2; Length 39;
Best Local Similarity 86.7%; Pred. No. 4.6e+02;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gcgggctcacagtgg 15
||||| |||||
Db 20 GCAGGCTGACAGTGG 34

RESULT 6
US-09-161-015-17/c
; Sequence 17, Application US/09161015A
; Patent No. 5965370
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF RHO EXPRESSION
; FILE REFERENCE: RTS-0015
; CURRENT APPLICATION NUMBER: US/09/161.015A
; CURRENT FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 17
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-161-015-17

Query Match 76.0%; Score 11.4; DB 2; Length 18;
Best Local Similarity 92.3%; Pred. No. 7.5e+02;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 14
||||| |||||
Db 17 CGGGCGCACAGTGG 5

RESULT 7
US-08-445-289B-25/c
; Sequence 25, Application US/08445289B
; Patent No. 5693467
; GENERAL INFORMATION:
; APPLICANT: Roblin III, Richard O.
; APPLICANT: Hu, Mendong
; APPLICANT: Tang, Jane S.
; APPLICANT: Lee, Sunmin
; TITLE OF INVENTION: A Mycoplasma PCR Testing System Using A
; TITLE OF INVENTION: Set of Mixed and Single Sequence Primers
; Patent No. 5693467
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Type Culture Collection
; STREET: 12301 Parklawn Drive
; CITY: Rockville
; STATE: Maryland
; COUNTRY: U.S.A.
; ZIP: 20852
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,289B
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Bade, Annette L.
; REGISTRATION NUMBER: 37,029
; REFERENCE/DOCKET NUMBER: RD00003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-231-5520
; TELEFAX: 310-816-4366
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-445-289B-25

Query Match 76.0%; Score 11.4; DB 1; Length 22;
Best Local Similarity 92.3%; Pred. No. 7.5e+02;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ggggtcacagtgg 15
||||| |||||
Db 13 GGGCACACAGTGG 1

RESULT 8
US-08-859-998-856/c
; Sequence 856, Application US/08859998
; Patent No. 5994076
; GENERAL INFORMATION:
; APPLICANT: Chenchik, Alex
; APPLICANT: Jokhadze, George
; APPLICANT: Bibilashvili, Robert
; TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1375
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: US
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,998
; FILING DATE: 21-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Field, Bret E.
; REGISTRATION NUMBER: 37,620
; REFERENCE/DOCKET NUMBER: 09096/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 856:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; OTHER INFORMATION: oligonucleotide primer
US-08-859-998-856

Query Match 73.3%; Score 11; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 ggtcacagtg 14
|||||
Db 11 GGCTCACAGT 1

RESULT 9
US-07-959-946-8/c
; Sequence 8, Application US/07959946
; Patent No. 5408038
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5408038th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/959,946
; FILING DATE: 19921008
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-959-946-8

Query Match 73.3%; Score 11; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ggtcacagtg 15
|||||
Db 19 GCTCACAGTG 9

RESULT 10
US-08-333-577-8/c
; Sequence 8, Application US/08333577
; Patent No. 5786206
; GENERAL INFORMATION:

; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; APPLICANT: Curtiss, Linda K.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 No. 5786206th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,577
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER: SCRF 234.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-333-577-8

Query Match 73.3%; Score 11; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ggtcacagtg 15
|||||
Db 19 GCTCACAGTG 9

RESULT 11
PCT-US92-08634-8/c
; Sequence 8, Application PC/TUS9208634
; GENERAL INFORMATION:
; APPLICANT: Smith, Richard K.
; APPLICANT: Koduri, Raju
; APPLICANT: Young, Stephen G.
; APPLICANT: Witztum, Joseph L.
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a
; TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
; ADDRESSEE: Milnamow, Ltd.
; STREET: 180 North Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08634
; FILING DATE: 19921009
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,706
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamson, Edward P.
; REGISTRATION NUMBER: 29,381
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)616-5400
; TELEFAX: (312)616-5460
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US92-08634-8

Query Match 73.3%; Score 11; DB 5; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 gctcacagtgg 15
| | | | | | | | | |
Db 19 GCTCACAGTGG 9

RESULT 12
US-09-166-203-9
; Sequence 9, Application US/09166203A
; Patent No. 5968826
; GENERAL INFORMATION:
; APPLICANT: Bennett, C. Frank
; APPLICANT: Condon, Tom P.
; APPLICANT: Cowsett, Lex M.
; TITLE OF INVENTION: ANTISENSE MODULATION OF INTEGRIN 4 EXPRESSION
; FILE REFERENCE: ISPH-0323
; CURRENT APPLICATION NUMBER: US/09/166,203A
; CURRENT FILING DATE: 1998-10-05
; NUMBER OF SEQ ID NOS: 60
; SEQ ID NO 9
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-166-203-9

Query Match 72.0%; Score 10.8; DB 2; Length 18;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gcgggctcacagt 14
| | | | | | | | | |
Db 1 gcggctctcagt 14

RESULT 13
US-08-206-176-10/c
; Sequence 10, Application US/08206176
; Patent No. 563940
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian

; APPLICANT: Dalrymple, Michael A
; APPLICANT: Prunkard, Donna E
; APPLICANT: Foster, Donald C
; TITLE OF INVENTION: Production of Fibrinogen in Transgenic
; TITLE OF INVENTION: Animals
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,176
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: BLGAMP4
; US-08-206-176-10

Query Match 72.0%; Score 10.8; DB 1; Length 24;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 cgggctcacagtgg 15
| | | | | | | | | |
Db 24 CTGGCTCACAGG 11

RESULT 14
US-08-371-001-2
; Sequence 2, Application US/08371001
; Patent No. 5783683
; GENERAL INFORMATION:
; APPLICANT: Morrison Ph.D., Richard
; TITLE OF INVENTION: Methods and Composition for Treating
; TITLE OF INVENTION: Tumor Cells
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gray Cary Ware & Freidenrich
; STREET: 401 "B" Street, Suite 1700
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/371,001
; FILING DATE: January 10, 1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brotman, Harris F.
REGISTRATION NUMBER: 35,461
REFERENCE/DOCKET NUMBER: P00095US0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 699-3630
TELEFAX: (619) 236-1048
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Other Nucleic Acid
US-08-371-001-2

Query Match 72.0%; Score 10.8; DB 1; Length 24;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 15
||| ||||| |||||
Db 1 CGAGCTCACTGTGG 14

RESULT 15
US-08-756-506-9/C
Sequence 9, Application US/08756506
Patent No. 5905185
GENERAL INFORMATION:
APPLICANT: Garner, Ian
APPLICANT: Cottingham, Ian R.
APPLICANT: Temperley, Simon M.
APPLICANT: Foster, Donald C.
APPLICANT: Sprecher, Cindy A.
APPLICANT: Prunkard, Donna E.
TITLE OF INVENTION: PROTEIN C PRODUCTION IN TRANSGENIC
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 1201 Eastlake Avenue East
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,506
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sawislak, Deborah A
REGISTRATION NUMBER: 37,438
REFERENCE/DOCKET NUMBER: 95-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6672
TELEFAX: 206-442-6678
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-756-506-9

Query Match 72.0%; Score 10.8; DB 2; Length 24;
Best Local Similarity 85.7%; Pred. No. 1.6e+03;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 cgggctcacagtgg 15
||| ||||| |||||
Db 24 CTGGCTCACAGAGG 11

Search completed: October 2, 2001, 16:03:54
Job time: 14598 sec

THIS PAGE BLANK (USPTO)